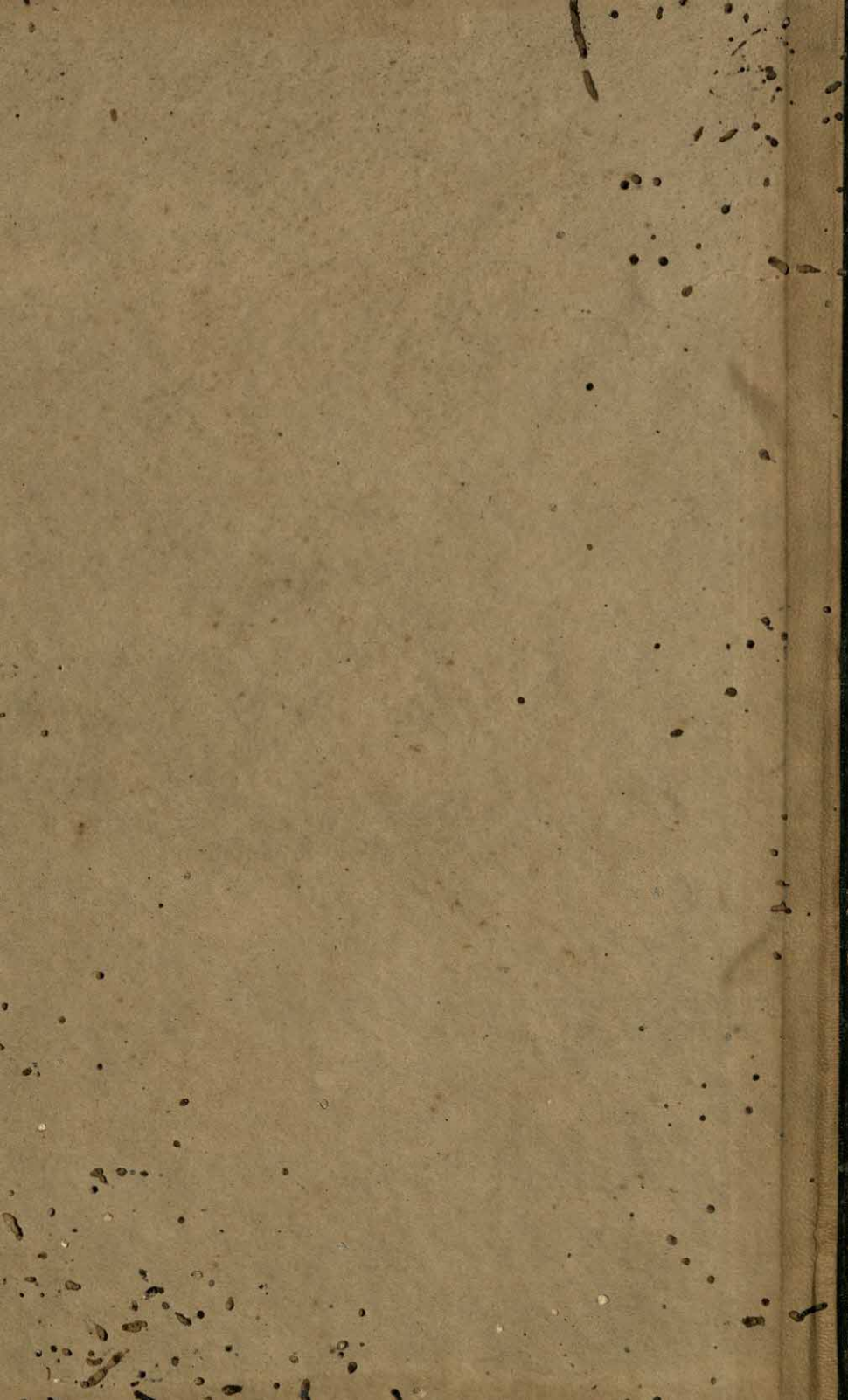
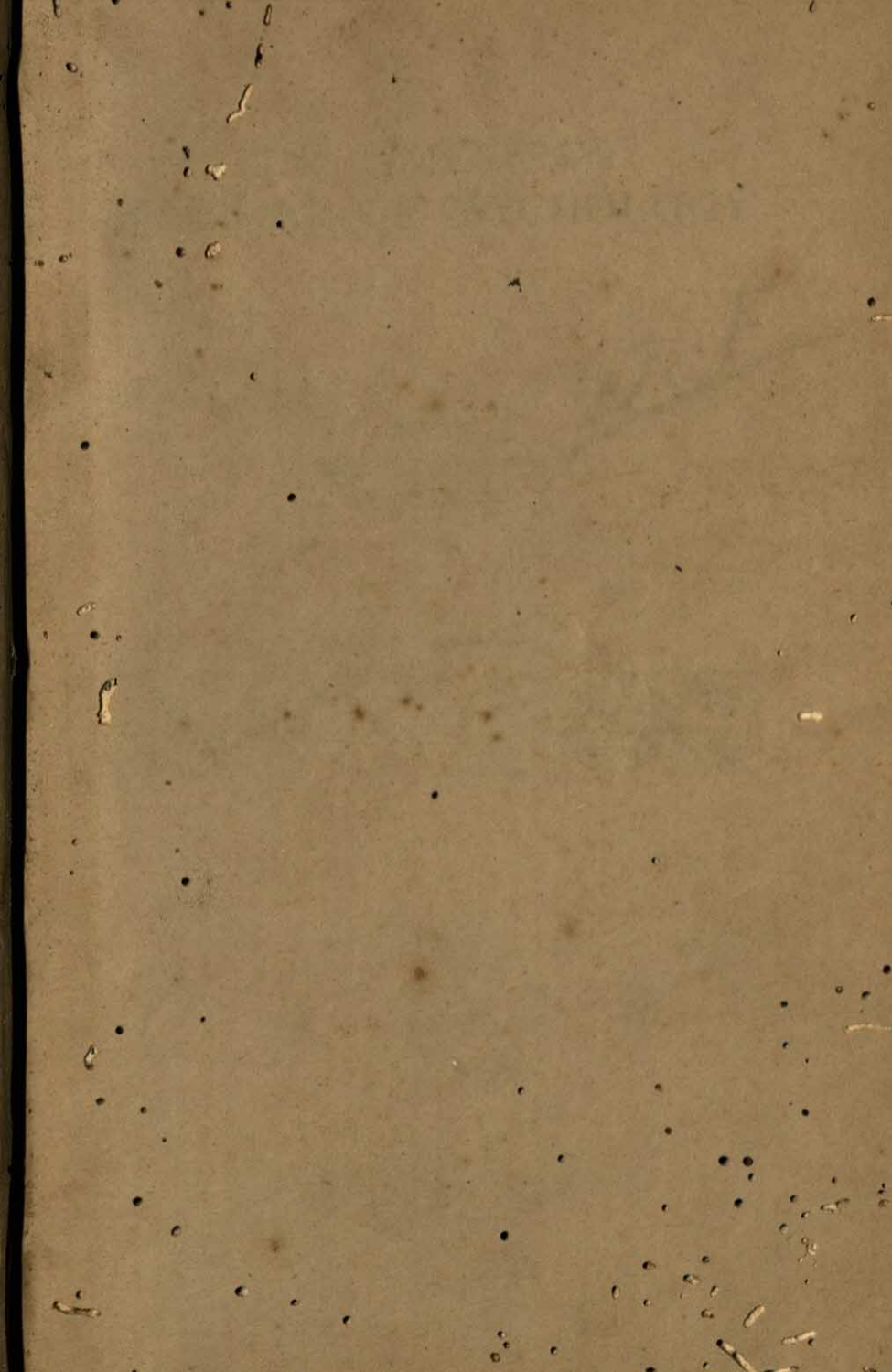
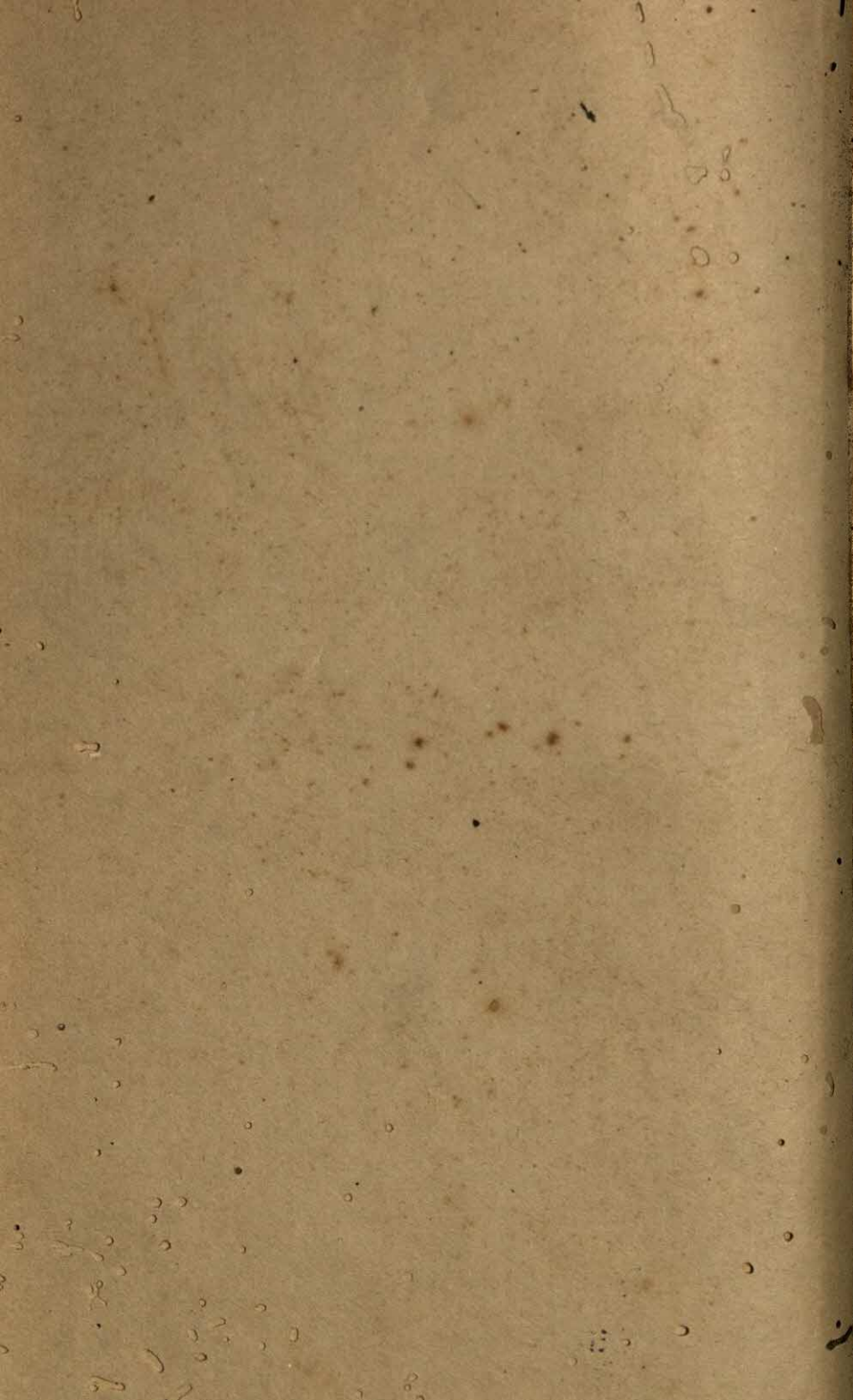




87

6.8.20





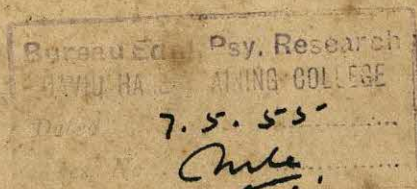
THE AMERICAN JOURNAL OF PSYCHOLOGY

EDITED BY

KARL M. DALLENBACH
University of Texas

WITH THE COÖPERATION OF

E. G. BORING, Harvard University; M. E. BITTERMAN, University of California; S. W. FERNBERGER, University of Pennsylvania; J. P. GUILFORD, University of Southern California; HARRY HELSON, University of Texas; E. R. HILGARD, Stanford University; W. S. HUNTER, Brown University; G. L. KREEZER, Washington University; D. G. MARQUIS, University of Michigan; R. M. OGDEN, Cornell University; W. B. PILLSBURY, University of Michigan



Vol. LXVII
MEZES HALL, THE UNIVERSITY OF TEXAS
AUSTIN, TEXAS
1954



7872
COPYRIGHT, 1955

BY

KARL M. DALLENBACH

Bureau Ednl. Psy. Research
DAVID H. ... KING COLLEGE
Dated ... 6. 8. 70
Access. No. ... 87

TABLE OF CONTENTS

ARTICLES AND NOTES

ADAMS, J. A., and REYNOLDS, B., Psychomotor Performance as a Function of Initial Level of Ability	268
ADAMS, P. A., The effect of Past Experience on the Perspective Reversal of a Tridimensional Figure	708
ADAMS, P. A., and POSTMAN, L., Performance Variables in the Experimental Analysis of the Law of Effect	612
ADAMS, P. A., and WALLACH, H., Binocular Rivalry of Achromatic Colors ..	513
AIKEN, E. G., and MOWRER, O. H., Contiguity vs. Drive-Reduction in Conditioned Fear: Temporal Variations in Conditioned and Unconditioned Stimulus	26
ANDREWS, T. G., COFER, C. N., and SHAPIRO, S., Transfer and Generalization of the Inhibitory Potential Developed in Rote Serial Learning	453
ANDREWS, T. G., LYBRAND, W. A., ROSS, S., Systemic Fatigue and Perceptual Organization	704
ATKINSON, R. L., and SALTZMAN, I. J., Comparisons of Incidental and Intentional Learning after Different Numbers of Stimulus Presentations	521
AUSTIN, P., and WALLACH, H., Recognition and the Localization of Visual Traces	338
BACHEM, A., Musical Memory Mechanics	360
BARTLEY, S. H., and HAHN, E. L., The Apparent Orientation of a Luminous Figure in Darkness	500
BEACH, F. A., The 1953 Meeting of Section I, AAAS	176
BECK, L. H., The Barnes-Czerny Apparatus for Demonstrating Quantal Fluctuations in Vision	152
BERGER, C., and MAHNEKE, A., Fatigue in Two Simple Visual Tasks	509
BILLINGSLEY, B. A., BITTERMAN, M. E., and FEDDERSEN, W. E., Discrimination Following Nondifferential Reinforcement with Differential Afferent Consequences	335
BINDRA, D., Thirteenth Annual Meeting of the Canadian Psychological Association	554
BITTERMAN, M. E., BILLINGSLEY, B. A., and FEDDERSEN, W. E., Discrimination Following Nondifferential Reinforcement with Differential Afferent Consequences	335
BITTERMAN, M. E., and BRIDGES, C. C., The Measurement of Autokinetic Movement	525
BITTERMAN, M. E., CRAWFORD, F. T., and MAYES, G. L., A Further Study of Differential Afferent Consequences in Nondifferential Reinforcement	717
BITTERMAN, M. E., DURYEA, R. A., and KRAUSKOPF, J., Threshold for Visual Form: Further Experiments	427
BITTERMAN, M. E., and ELAM, C. B., Discrimination Following Varying Amounts of Nondifferential Reinforcement	135
BITTERMAN, M. E., HOCHBERG, J. E., and KRAUSKOPF, J., Threshold for Visual Form: A Diffusion Model	206
BITTERMAN, M. E., and MCCONNELL, J. V., The Rôle of Set in Successive Discrimination	129
BORING, E. G., The Nature and History of Experimental Control	573
BORING, E. G., The 1954 Meeting of the American Philosophical Society	370
BORING, E. G., The 1954 Meeting of the National Academy of Sciences	369
BOUSEFIELD, W. A., COHEN, B. H., and SEDGEWICK, C. H. W., Certain Temporal Characteristics of the Recall of Verbal Associates	111
BRIDGES, C. C., and BITTERMAN, M. E., The Measurement of Autokinetic Movement	525
BRIGGS, G. E., BROGDEN, W. J., and THOMPSON, R. F., The Effect of Angle	

of Tilt upon the Trigonometric Relationship of Precision and Angle of Linear Pursuit-Movements	475
BROGDEN, W. J., BRIGGS, G. E., and THOMPSON, R. F., The Effect of Angle of Tilt upon the Trigonometric Relationship of Precision and Angle of Linear Pursuit-Movements	475
BROWN, K. T., An Optical Illusion of Spontaneous Fluctuations in Apparent Size	553
BROWN, R. H., and CONKLIN, J. E., The Lower Threshold of Visible Movement as a Function of Exposure-Time	104
BROŽEK, J., and TAYLOR, H. L., Tests and Motor Functions in Investigations on Fitness	590
BURNHAM, R. W., Binocular Subjective Colors and the Visual Mechanism ..	492
CARMICHAEL, L., Walter Samuel Hunter: 1889-1954	732
CHRISTMAN, R. J., Shifts in Pitch as a Function of Prolonged Stimulation with Pure Tones	484
COFER, C. N., ANDREWS, T. G., and SHAPIRO, S., Transfer and Generalization of the Inhibitory Potential Developed in Rote Serial Learning	453
COHEN, B. H., BOUSFIELD, W. A., and SEDGEWICK, C. H. W., Certain Temporal Characteristics of the Recall of Verbal Associates	111
COHN, V. H., and GROSS, N. B., The Effect of Vitamin-B Deficiency on the Hoarding Behavior of Rats	124
COLGAN, C. M., Critical Flicker Frequency, Age, and Intelligence	711
CONKLIN, J. E., Senders on Response-Sequences	363
CONKLIN, J. E., and BROWN, R. H., The Lower Threshold of Visible Movement as a Function of Exposure-Time	104
CRAWFORD, F. T., BITTERMAN, M. E., and MAYES, G. L., A Further Study of Differential Afferent Consequences in Nondifferential Reinforcement	717
CRONBACH, L. J., Twenty-Sixth Annual Meeting of the Midwestern Psychological Association	368
CRUMBAUGH, J. C., Temporal Changes in the Memory of Visually Perceived Form	647
DALLENBACH, K. M., An Acknowledgment	734
DALLENBACH, K. M., An Editorial Announcement	571
DALLENBACH, K. M., Errata	178
DAVIS, R. C., SIDDONS, G. F., and STOUT, G. L., Apparatus for Recording Autonomic States and Changes	343
DAY, R. H., The Effect of Size of Target on Accuracy of Aim	659
DEATHERAGE, B. H., Figural After-Effects in Stroboscopic Movement	259
DEESE, J., and HARDMAN, G. W., An Analysis of Errors in Retroactive Inhibition of Rote Verbal Learning	299
DINGMAN, H. F., and GUILFORD, J. P., A Validation Study of Ratio-Judgment Methods	395
DURYEA, R. A., BITTERMAN, M. E., and KRAUSKOPF, J., Threshold for Visual Form: Further Experiments	427
EDWARDS, W., Methods for Computing Uncertainties	164
EDWARDS, W., Probability-Preferences Among Bets with Differing Expected Values	56
EDWARDS, W., The Reliability of Probability-Preferences	68
EDWARDS, W., Variance Preferences in Gambling	441
ELAM, C. B., and BITTERMAN, M. E., Discrimination Following Varying Amounts of Nondifferential Reinforcement	133
FEDDERSEN, W. E., BILLINGSLEY, B. A., and BITTERMAN, M. E., Discrimination Following Nondifferential Reinforcement with Differential Afferent Consequences	335
TERGUON, G. A., Twelfth Annual Meeting of the Canadian Psychological Association	178
FERNBERGER, S. W., The Prestige and Impact of Various Psychologists on Psychology in America	288

GHENT, L., SEMMES, J., TEUBER, H. L., and WEINSTEIN, S., Performance on Complex Tactual Tasks after Brain Injury in Man: Analyses by Locus of Lesion	220
GREEN, E. J., An Anchoring Effect in the Operant Responding of Rats	141
GROSS, N. B., and COHN, V. H., The Effect of Vitamin-B Deficiency on the Hoarding Behavior of Rats	124
GRUBER, H. E., The Relation of Perceived Size to Perceived Distance	411
GUILFORD, J. P., System in the Relationship of Affective Value to Frequency and Intensity of Auditory Stimuli	691
GUILFORD, J. P., and DINGMAN, H. F., A Validation Study of Ratio-Judgment Methods	395
HAGGARD, E. A., and THOMPSON, W. R., Electrodes for Measuring the GSR in Small Animals	357
HAHN, E. L., and BARTLEY, S. H., The Apparent Orientation of a Luminous Figure in Darkness	500
HALL, J. F., Learning as a Function of Word-Frequency	138
HANAWALT, N. G., Recurrent Images: New Instances and a Summary of the Older Ones	170
HARDMAN, G. W., and DEESE, J., An Analysis of Errors in Retroactive Inhibition of Rote Verbal Learning	299
HARRISON, J. M., and MAYER, C. E., An Inexpensive Square Wave Stimulator	353
HARTMAN, E. B., The Influence of Practice and Pitch-Distance Between Tones on the Absolute Identification of Pitch	1
HELSON, H., and MICHELS, W. C., A Quantitative Theory of Time-Order Effects	327
HELSON, H., and MICHELS, W. C., A Reconciliation of the VEG Scale with Fechner's Law	677
HELSON, H., MICHELS, W. C., and STURGEON, A., The Use of Comparative Rating Scales for the Evaluation of Psychophysical Data	321
HOCHBERG, J. E., BITTERMAN, M. E., and KRAUSKOPF, J., Threshold for Visual Form: A Diffusion Model	205
HOFFMAN, E. L., and ROHRER, J. H., An Apparatus for Studying the Perception of Light-Movement	143
HOWES, D., and OSGOOD, C. E., On the Combination of Associative Probabilities in Linguistic Contexts	241
JAFFE, R., Kinesthetic After-Effects Following Cerebral Lesions	668
JENKINS, J. J., and KING-ELLISON, P., The Deviational Threshold of Visual Recognition as a Function of Word-Frequency	700
JOHNSON, H. M., On Verifying Hypotheses by Verifying Their Implicates ..	723
JONES, F. N., An Olfactory Permitting Stimulus Specification in Molar Terms ..	147
JONES, M. H., A Study of the 'Common Chemical Sense'	696
KARLIN, L., An Improved Circuit for Connecting the Hunter Timer and Similar Timers in Sequence	358
KENNEDY, J. L., Fiftieth Annual Meeting of the Society of Experimental Psychologists	367
KING-ELLISON, P., and JENKINS, J. J., The Durational Threshold of Visual Recognition as a Function of Word-Frequency	700
KRAUSKOPF, J., Figural After-Effects in Auditory Space	278
KRAUSKOPF, J., The Magnitude of Figural After-Effects as a Function of the Duration of the Test-Period	684
KRAUSKOPF, J., BITTERMAN, M. E., and DURYEA, R. A., Thresholds for Visual Form: Further Experiments	427
KRAUSKOPF, J., BITTERMAN, M. E., and HOCHBERG, J. E., Threshold for Visual Form: A Diffusion Model	205
LACEY, O. L., Forty-Sixth Annual Meeting of the Southern Society for Philosophy and Psychology	369
LANE, G., Twenty-Fifth Annual Meeting of the Eastern Psychological Association	365

LAW, O. T., Preference in the Rat for Vertical or Horizontal Stripes After Training on a White-Black Discrimination	714
LEIBOWITZ, H., The Use and Calibration of the 'Maxwellian View' in Visual Instrumentation	530
LOCKE, N., MULL, H. K., and ORD, N., The Effect of Two Brightness Factors upon the Rate of Fluctuation of Reversible Perspectives	341
LYBRAND, W. A., ANDREWS, T. G., and ROSS, S., Systemic Fatigue and Perceptual Organization	704
MACCASLIN, E. F., Successive and Simultaneous Discrimination as a Function of Stimulus-Similarity	308
MAHNEKE, A., and BERGER, C., Fatigue in Two Simple Visual Tasks.	509
MAYER, C. E., and HARRISON, J. M., An Inexpensive Square Wave Stimulator	353
MAYES, G. L., BITTERMAN, M. E., and CRAWFORD, F. T., A Further Study of Differential Afferent Consequences in Nondifferential Reinforcement	717
MCCONNELL, J. V., and BITTERMAN, M. E., The Role of Set in Successive Discrimination	129
MEYER, M. F., A Test of the Hydraulic Theory of Cochlear Mechanics Through Multiple Frequency Beats	39
MEYER, M. F., Auditory Fatigue Beyond and Within the Compass of the Voice	538
MICHEL, W. C., and HELSON, H., A Quantitative Theory of Time-Order Effects	327
MICHEL, W. C., and HELSON, H., A Reconciliation of the VEG Scale with Fechner's Law	677
MICHEL, W. C., HELSON, H., and STURGEON, A., The Use of Comparative Rating Scales for the Evaluation of Psychophysical Data	321
MOWRER, O. H., and AIKEN, E. G., Contiguity vs. Drive-Reduction in Conditioned Fear: Temporal Variations in Conditioned and Unconditioned Stimulus	26
MOWRER, O. H., and SOLOMON, L. N., Contiguity vs. Drive-Reduction in Conditioned Fear: The Proximity and Abruptness of Drive-Reduction	15
MULHOLLAND, T. B., Some Remarks Concerning the Use of the Method of Adjustment with Implicit Standards	546
MULL, H. K., LOCKE, N., and ORD, N., The Effect of Two Brightness Factors upon the Rate of Fluctuation of Reversible Perspectives	341
NAFE, J. P., Harry Miles Johnson: 1885-1953	179
NEWHALL, S. M., Comparability of the Method of Single Stimuli and the Method of Paired Comparisons	96
ORD, N., LOCKE, N., and MULL, H. K., The Effect of Two Brightness Factors upon the Rate of Fluctuation of Reversible Perspectives	341
OSGOOD, C. E., and HOWES, D., On the Combination of Associative Probabilities in Linguistic Contexts	241
PAGE, H. A., and WYCKOFF, L. B., A Grid for Administering Shock	154
POSTMAN, L., and ADAMS, P. A., Performance Variables in the Experimental Analysis of the Law of Effect	612
POSTMAN, L., and TUMA, A. H., Latent Learning in Human Subjects	119
PRATT, C. C., The XIV International Congress of Psychology	551
PRENTICE, W. C. H., Red Light and the Function of the Rods	727
PRENTICE, W. C. H., Visual Recognition of Verbally Labeled Figures	315
REYNOLDS, B., and ADAMS, J. A., Psychomotor Performance as a Function of Initial Level of Ability	268
ROBINSON, E. J., The Influence of Photometric Brightness on Judgments of Size	464
ROGERS, L. S., Twenty-Fourth Annual Meeting of the Rocky Mountain Branch of the American Psychological Association	366
ROHRER, J. H., and HOFFMAN, E. L., An Apparatus for Studying the Perception of Light-Movement	143
ROSS, S., ANDREWS, T. G., and LYBRAND, W. A., Systemic Fatigue and Perceptual Organization	704

SALTZMAN, I. J., and ATKINSON, R. L., Comparisons of Incidental and Intentional Learning after Different Numbers of Stimulus Presentations	521
SANFORD, F. H., Sixty-First Annual Meeting of the American Psychological Association	177
SANFORD, F. H., Sixty-Second Annual Meeting of the American Psychological Association	730
SEDGEWICK, C. H. W., BOUSFIELD, W. A., and COHEN, B. H., Certain Temporal Characteristics of the Recall of Verbal Associates	111
SEMMES, J., GHENT, L., WEINSTEIN, S., and TEUBER, H. L., Performance on Complex Tasks After Brain Injury in Man: Analyses by Locus of Lesion	220
SHAPIRO, S., ANDREWS, T. G., and COFER, C. N., Transfer and Generalization of the Inhibitory Potential Developed in Rote Serial Learning	453
SIDDONS, G. F., DAVIS, R. C., and STOUT, G. L., Apparatus for Recording Autonomic States and Changes	343
SMITH, K., 'Attraction' in Figural After-Effects	174
SOLOMON, L. N., and MOWRER, O. H., Contiguity vs. Drive-Reduction in Conditioned Fear: The Proximity and Abruptness of Drive-Reduction	15
STOUT, G. L., DAVIS, R. C., and SIDDONS, G. F., Apparatus for Recording Autonomic States and Changes	343
STURGEON, A., HELSON, H., and MICHELS, W. C., The Use of Comparative Rating Scales for the Evaluation of Psychophysical Data	321
TAYLOR, H. L., and BROŽEK, J., Tests of Motor Functions in Investigations on Fitness	590
TEUBER, H. L., GHENT, L., SEMMES, J., and WEINSTEIN, S., Performance on Complex Tactual Tasks After Brain Injury in Man: Analyses by Locus of Lesion	220
THOMAS, G. J., The Effects on Critical Flicker Frequency of Interocular Differences in Intensity and in Phase Relations of Flashes of Light	682
THOMPSON, R. F., BRIGGS, G. E., and BROGDEN, W. J., The Effect of Angle of Tilt upon the Trigonometric Relationship of Precision and Angle of Linear Pursuit-Movements	475
THOMPSON, W. R., and HAGGARD, E. A., Electrodes for Measuring the GSR in Small Animals	357
TUMA, A. H., and POSTMAN, L., Latent Learning in Human Subjects	119
TYLER, L. E., Thirty-Fourth Annual Meeting of the Western Psychological Association	553
VINCENT, D. F., The Earliest Formulae Used in Factor Analysis	155
WALLACH, H., and ADAMS, P. A., Binocular Rivalry of Achromatic Colors	513
WALLACH, H., and AUSTIN, P., Recognition and the Localization of Visual Traces	338
WARREN, J. M., Reversal Discrimination as a Function of the Number of Reinforcements During Pre-Training	720
WARREN, J. M., An Analysis of the Formation of Visual Discriminative Habits of Rhesus Monkeys	517
WEINSTEIN, S., GHENT, L., SEMMES, J., and TEUBER, H. L., Performance on Complex Tactual Tasks After Brain Injury in Man: Analyses by Locus of Lesion	220
WERTHEIMER, M., Constant Errors in the Measurement of Figural After-Effects	543
WYCKOFF, L. B., and PAGE, H. A., A Grid for Administering Shpck	154

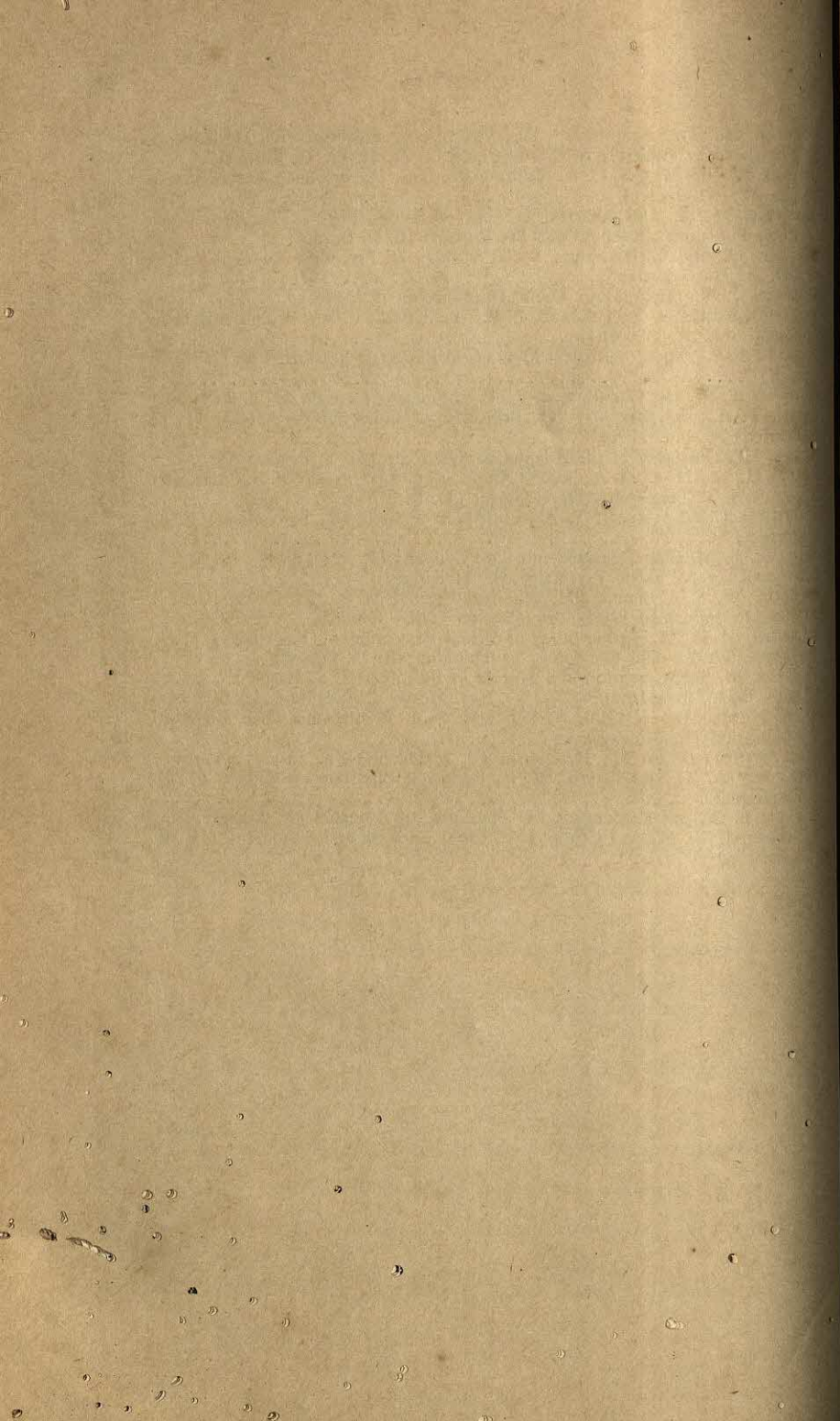
BOOK REVIEWS

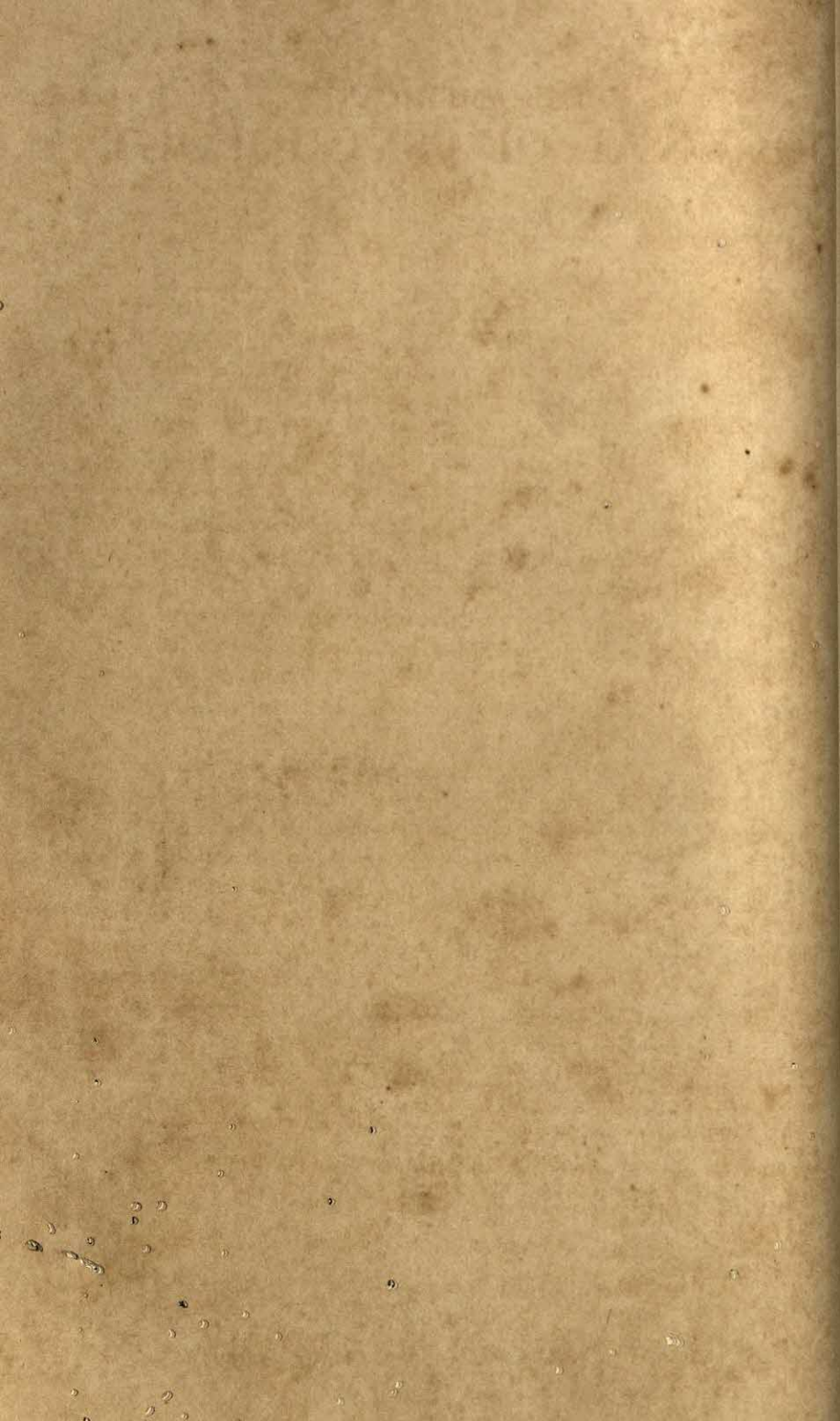
(The reviewer's name appears in parentheses after the title of the work.)

ABBOTT, B. C., and PIRENNE, M. H. Translators, Piéron's The Sensations: Their Functions, Processes and Mechanisms (E. G. Boring)	383
ALLPORT, G. W., The Nature of Prejudice (R. Christie)	742
ANDERSON, I. H., and DEARBORN, W. F., The Psychology of Teaching Reading (M. A. Tinker)	197

ASCH, S. E., Social Psychology (R. B. MacLeod)	371
ASHBY, W. R., Design for a Brain (J. Krauskopf)	391
BELLAK, L., Manic Depressive Psychosis (F. J. Moore)	393
BRUNSWIK, E., The Conceptual Framework of Psychology (J. E. Hochberg)	386
CARROLL, J. B., The Study of Language (R. M. Ogden)	568
CHILD, I. L., and WHITING, J. W. M., Child Training and Personality (H. W. Stevenson)	749
COHEN, L. H., Murder, Madness and the Law (H. P. Weld)	390
CONANT, J. B., Modern Science and Modern Man (E. G. Boring)	192
DEARBORN, W. F., and ANDERSON, I. H., The Psychology of Teaching Reading (M. A. Tinker)	197
DOLL, E. A., The Measurement of Social Competence: A Measure for the Vineland Maturity Scale (I. Iscoe)	754
FENLASON, A. F., Essentials in Interviewing (P. C. Smith)	394
FERGUSON, L. W., Personality Measurement (J. P. Guilford)	195
FLOYD, W. F., and WELFORD, A. T., Editors, Symposium on Fatigue (J. M. Vanderplas)	567
GARDNER, M., In the Name of Science (E. G. Boring)	204
GELDARD, F. A., The Human Senses (E. G. Boring)	383
GILLIN, J., Editor, For A Science of Social Man (F. K. Berrien)	748
GILMER, B. v. H., and KARN, H. W., Readings in Industrial and Business Psychology (J. S. Gray)	392
GRINKER, R. R., and ROBBINS, F. P., Psychosomatic Case Book (T. R. Sarbin)	754
GUTHRIE, E. F., The Psychology of Learning (M. E. Bitterman)	394
HILGARD, E. R., Introduction to Psychology (D. Krech)	561
HILGARD, E. R., PUMPIAN-MINDLIN, E., and KUBIE, L. S., Psychoanalysis As Science (R. R. Blake)	393
HIRSH, I. J., The Measurement of Hearing (L. A. Jeffress)	751
HOFSTÄTTER, P. R., Einführung in die Sozialpsychologie (H. J. Eysenck)	753
HORAVA, A., and SELYE, H., Second Annual Report on Stress (M. E. Bitterman)	394
HULL, C. L., A Behavior System: An Introduction to Behavior Theory Concerning the Individual Organism (R. Leeper)	375
KARN, H. W., and GILMER, B. von H., Readings in Industrial and Business Psychology (J. S. Gray)	392
KECKSEMETI, P., Meaning, Communication, and Value (J. B. Carroll)	744
KENTUCKY SYMPOSIUM, THE, Learning Theory, Personality Theory, and Clinical Research (M. E. Bitterman)	743
KILPATRICK, F. P. Editor, Human Behavior from the Transactional Point of View (N. Pastore)	379
KLINEBERG, O., Social Psychology (R. B. MacLeod)	752
KUBIE, L. S., HILGARD, E. R., and PUMPIAN-MINDLIN, E., Psychoanalysis As Science (R. R. Blake)	393
LAMB, G., Translator, Nuttin's Psychoanalysis and Personality (L. W. Crafts)	388
LAWRENCE, M., and WEVER, E. G., Physiological Acoustics (K. Lowy)	735
LE CRON, L. M., Editor, Experimental Hypnosis (F. L. Marcuse)	199
MAIER, N. R. F., Principles of Human Relations (R. S. Harper)	202
MATHEWS, R. W., and WALLS, G. L., New Means of Studying Color Blindness and Normal Foveal Color Vision, With Some Results and Their Genetical Implication (E. Murray)	182
MENNINGER, K. A., A Manual for Psychiatric Case Study (P. Worchel)	391
MYERS, C. A., and PIGORS, P., Editors, Readings in Personnel Administration (R. W. Kleemeier)	200
NUTTIN, J., Psychoanalysis and Personality. Translated by LAMB, G. (L. W. Crafts)	388
OSGOOD, C. E., Method and Theory in Experimental Psychology (W. C. H. Prentice)	555
OVERHOLSER, W., The Psychiatrist and the Law (H. P. Weld)	747
PICKFORD, R. W., Individual Differences in Colour Vision (H. Helson)	189

PIÉRON, H., The Sensations: Their Functions, Processes and Mechanisms, Translated by ABBOTT, B. C. and PIRENNE, M. H. (E. G. Boring)	383
PIGORS, P., and MYERS, C. A., Editors, Readings in Personnel Administration (R. W. Kleemeier)	200
PIRENNE, M. H., and ABBOTT, B. C., Translators, Piéron's The Sensations: Their Functions, Processes and Mechanisms (E. G. Boring)	383
POLLAK, O., and Collaborators, Social Science and Psychotherapy for Children (D. B. Harris)	198
POWELL, J. N., Anatomy of Public Opinion (G. B. Vetter)	190
PUMPIAN-MINDLIN, E., HILGARD, E. R., and KUBIE, L. W., Psychoanalysis As Science (R. R. Blake)	393
RAO, C. R., Advanced Statistical Methods in Biometric Research (W. H. Holtz- man)	566
ROBBINS, F. P., and GRINKER, R. R., Psychosomatic Case Book (T. R. Sarbin)	754
RYAN, T. A., and SMITH, P. C., Principles of Industrial Psychology (F. K. Berrien)	739
SELYE, H., The Story of the Adaptation Syndrome (M. E. Bitterman)	393
SELYE, H., and HORAVA, A., Second Annual Report on Stress (M. E. Bitterman)	394
SLOTKIN, J. S., Personality Development (D. R. Brown)	203
SMITH, P. C., and RYAN, T. A., Principles of Industrial Psychology (F. K. Berrien)	739
SPINLEY, B. M., The Deprived and the Privileged (R. S. Harper)	570
STRECKER, E. A., Basic Psychiatry (F. McKinney)	392
TAYLOR, W. S., Dynamic and Abnormal Psychology (L. W. Crafts)	746
TIFFIN, J., Industrial Psychology (B. von Haller Gilmer)	202
VERNON, M. D., A Further Study of Visual Perception (J. Krauskopf)	204
VINACKE, W. E., The Psychology of Thinking (R. M. Ogden)	194
VITELES, M. S., Motivation and Morale in Industry (R. S. Harper)	751
WALLS, G. L., and MATHEWS, R. W., New Means of Studying Color Blindness and Normal Foveal Color Vision, with Some Results and Their Genetical Implication (E. Murray)	182
WEITZENHOFFER, A. M., Hypnotism (F. L. Marcuse)	389
WELFORD, A. T., and FLOYD, W. F., Editors, Symposium on Fatigue (J. M. Vanderplas)	567
WEVER, E. G., and LAWRENCE, M., Physiological Acoustics (K. Lowy)	735
WHITING, J. W. M., and CHILD, I. L., Child Training and Personality (H. W. Stevenson)	749





THE AMERICAN JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXVII

MARCH, 1954

No. 1

THE INFLUENCE OF PRACTICE AND PITCH-DISTANCE BETWEEN TONES ON THE ABSOLUTE IDENTIFICATION OF PITCH

By E. B. HARTMAN, U.S.N. Submarine Base, New London, Connecticut

The term *absolute pitch* customarily implies an ability to name isolated musical tones—an ability possessed by few persons. The usual experimental test of this facility requires *O* to identify by name (C, G#, F, etc.) a series of tones presented singly, each being separated from the one preceding by a designated time-interval. If *O*'s average error of judgment satisfies a certain criterion (musical semitone or less), he is said to have absolute pitch; if his error exceeds the maximum allowed, he is not credited with absolute pitch.

Certain writers, Gough,¹ Köhler,² Meyer,³ and Mull,⁴ have attempted to show that *Os* who fail at first to meet the standard can, nevertheless, do so after training. The results of these studies, however, are largely equivocal, due to failure to control such factors as the number of tones in the series, the complexity of the stimuli, interstimulus time-interval, and pitch-distance between tones. It appears that what is needed is a careful determination of the effects of those factors upon absolute judgments of tones that are presented serially. How many tones can the average *O* adequately consider by the method of absolute judgment? How large must the pitch-

*Accepted for publication January 15, 1953. This paper is a portion of a dissertation submitted to the Faculty of the Graduate School of the University of Connecticut in partial fulfillment of the requirements for the degree of Master of Arts. The research was supervised by Professor Alvin M. Liberman and Dr. J. Donald Harris.

¹Evelyn Gough, The effects of practice on judgments of absolute pitch, *Arch. Psychol.*, 7, 1922 (No. 47), 1-93.

²Wolfgang Köhler, Akustische untersuchungen III, *Zsch. f. Psychol.*, 72, 1915, 159-177.

³Max Meyer, Is the memory of absolute pitch capable of development by training? *Psychol. Rev.*, 6, 1899, 514-516.

⁴H. K. Mull, The acquisition of absolute pitch, this JOURNAL, 36, 1925, 469-495.

distance be between successive tones before *O* can make correct identifications? How are absolute judgments affected by increasing the number of tones or by decreasing the pitch-distance between tones? These questions, besides bearing on the issue of absolute pitch, are of general interest with regard to broader problems of discrimination. We should like to know, for example, how many different pitch-qualities the average *O* can identify absolutely for purposes of practical communication.

In the interest of extending our knowledge about the rôle of pitch in absolute judgments, this paper poses the following questions: How widely must a given number of pure tones be separated in pitch before the average *O* can make correct absolute judgments? What happens to the accuracy of absolute judgments of pitch when *O* is presented smaller and smaller pitch separations? How is the amount of information which *O* receives about the series related to the pitch-distance between tones? Are the effects of pitch-separation on absolute judgments significantly influenced by practice? How does disuse affect the ability?

APPARATUS AND PROCEDURE

Four series of nine pure tones, separated by either 50, 100, 200, or 300 mels, were recorded at a constant intensity. In recording each separate series, the tones were placed discretely on the surface of the record to permit individual presentation. Each tone was of 2-sec. duration and was preceded by a 6-sec. burst of white noise. A General Radio Company Interpolation Oscillator, Type 617-C, was used to generate the tones. In reproduction, an attenuator situated between the phonograph and Permoflux PDR-8 headphones allowed *E* to control the intensity. Since a soundproof room was not available, the experiment took place in a reasonably quiet room. *O* sat facing a chart bearing a column of circles numbered from 1 to 9.

Observers. Twelve *O*s with normal hearing were separated into four groups on the basis of their ability to make absolute identifications of 36 pure semitones between *C*₆ and *B*₅. Each group yielded an average semitone error deviating either side of 4.0 semitones by less than a quartertone.

Instructions. The following instructions were read to *O*:

This is a test to see how well you can identify tones. There will be nine tones, which you are to identify by number. Listen while I sound each tone and give you its number. Notice that tone number one is the lowest, and tone number nine the highest. Notice also that each tone is preceded by a burst of noise. (*E* sounded each tone once, beginning with the lowest, and called each by number.)

Now turn to your answer sheet and find Item 1. I am going to sound one of the tones. You are to decide which tone it is, then write its number opposite Item 1. You may find the chart helpful in making your decision. Notice that there are nine circles numbered one to nine, or low to high. They will help you picture the position of the tones.

Do not take a long time answering, for I wish to correct your judgment before giving you the next tone. Try not to think of the previous tone while waiting for the next. Since a burst of noise precedes each tone, you might regard the noise as a 'ready' signal.

O was then given the following sequence, which was repeated every 30 sec.: a burst of noise, given to 'wash-out' the perseverative effects of preceding stimuli (6-sec.), an interval of silence (1-sec), a pure tone (2-sec). The tones were presented in a predetermined order, designed in such a way that each tone followed every other tone an equal number of times. Loudness level was about 50 phons. *E* attenuated each tone to approximately equal-loudness with the help of an equal-loudness chart based on two normal *Os*. When *O* had identified a particular stimulus, *E* told which stimulus had actually been given.

There were two experimental sessions every week for eight test-weeks; each session consisting of 72 judgments, or 8 judgments per stimulus. During Session 10 (Week 5), *O* was not told which stimuli had been presented, in order that *E* might determine how much he was relying on correction. Similarly, Sessions 15 and 16 (Week 8) were uncorrected, to make possible an estimate of unaided level of performance. Two months after the completion of the main series, *O* was given two additional sessions to establish how much he had forgotten.

• RESULTS AND DISCUSSION

Effects of pitch-separation. The customary treatment of the data from judgments of absolute pitch has been in terms of average error. The family of curves in Fig 1 expresses average error in mels against the week of practice, preserving the mel distance between tones as the parameter. These curves answer the question: By what pitch-distance, on the average, do the several experimental groups fail to identify the stimulus?

It will be noted (1) that no mel-group achieves 100% accuracy. This is perhaps understandable, inasmuch as perfect performance requires a total of 432 correct judgments per week—a rigorous criterion.

(2) In general, the smaller the pitch distance between tones, the smaller is the average error of judgment. The 50 and 100 mel-groups not only begin the experiment making smaller errors than do the larger mel groups, but they are able to maintain this advantage throughout the entire course of training.

(3) Learning curves for the 200 and 300 mel-groups drop sharply to asymptotes in about four weeks, while those for the 50 and 100 mel-groups show a more gradual descent, along with indications that learning is still occurring even at the end of training.

Gough, after training *Os* extensively to make absolute judgments of semitones on the piano, describes her learning curves as showing an irregular but gradual improvement, with no sign of a plateau at any point.⁵ Wedell's *Os*, on the contrary, judged tones separated by roughly $2\frac{1}{2}$ semitones and yielded curves which dropped quickly to asymptotes.⁶ The discrepancy between these two reports is accounted for

⁵ Gough, *op. cit.*, 86.

⁶ C. H. Wedell, The nature of the absolute judgment of pitch, *J. Exper. Psychol.* 17, 1934, 494 f.

by the present findings, showing that with narrow pitch-separations improvement occurs slowly and is incomplete even after considerable training, while with wider pitch-separations improvement is rapid and complete within relatively few practice sessions.

(4) Abandoning correction causes the average error to increase for all groups. For all but the 200 mel-group, this increase varies directly with

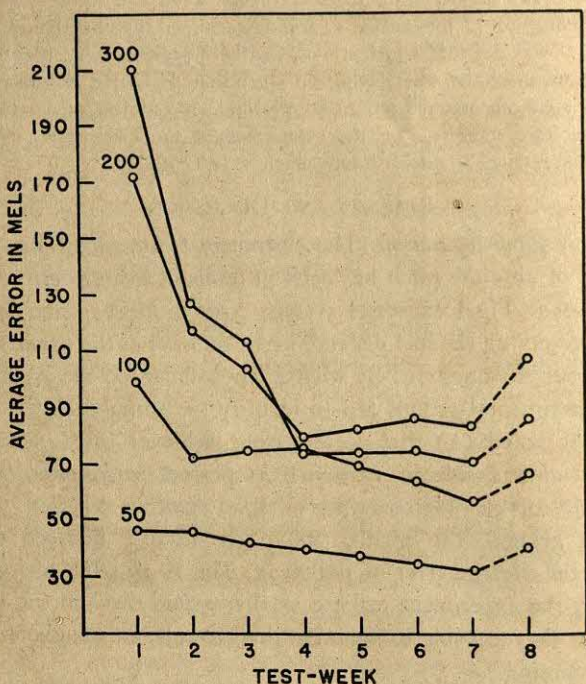


FIG. 1. ACCURACY OF ABSOLUTE JUDGMENTS OF NINE PURE TONES DURING EIGHT TEST-WEEKS

The values for Week 5 are interpolated, since one of the two sessions for that week was uncorrected and the results of the two sessions could not be averaged. The broken lines at right indicate the increase in error occurring when correction is withheld. Parameter is interstimulus pitch-distance in mels.

the size of the pitch-distance between tones. The anomaly is perhaps understandable in terms of the small number of Os employed.

(5) The wider the pitch-distance between tones, the more the average error is reduced by practice. Apparently, the degree of benefit derived from practice is not simply a function of the amount of time spent practicing, but varies also with the pitch-separation of tones in the series. This relation

is shown more clearly in Fig. 2, which plots the absolute reduction in average error occurring over seven weeks against the pitch-separation of tones in mels. The result is seen to be an approximate straightline function over the period of testing covered. This tendency for benefit from practice to increase with the pitch-separation, together with the fact that the average errors of the 200 and 300 mel-groups at the beginning of training are more nearly alike than are those of any two other mel-groups, explains how the 300 mel-group is able to surpass the performance of the 200 mel-group,

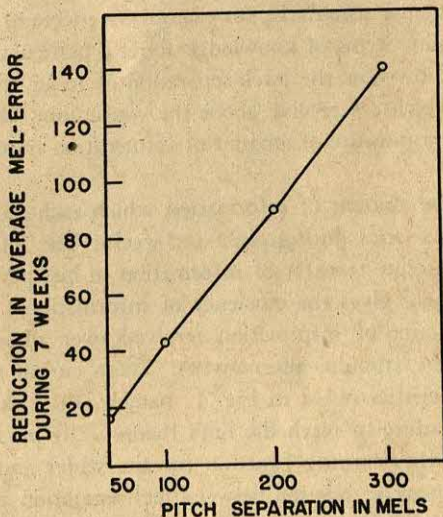


FIG. 2. EFFECTS OF SEVEN WEEKS' PRACTICE IN IDENTIFYING NINE PURE TONES AS A FUNCTION OF THE PITCH-DISTANCE BETWEEN NEIGHBORING TONES

even though the former's coarse separations entail errors of greater magnitude.

The fact that the average mel-error decreases with the pitch-separation cannot be taken to mean that *O*'s discrimination improves as the pitch-distance between tones becomes smaller and smaller. Consider the result of asking the *O*s to make absolute judgments of nine tones separated by infinitesimal frequency steps. Though the *O*s could perform no better than chance, their average errors, besides being of small absolute magnitude, would tend to be smaller than would those of a much coarser pitch-group committing occasional errors. Since all mel-groups in the present study were operating short of perfect identification, those groups judging the coarser pitch-separations were disproportionately penalized for guessing.

This fact becomes clearer when the percentage of correct judgments made by the various groups is taken into account, in which case it is found that the smaller the pitch-separation, the less often is the stimulus correctly identified. This is in agreement with Mull's finding that a reduction in the pitch-distance lowers both the number of correct identifications and the average error of judgment.⁷

Effects of pitch separation. Actually, the average error of *O*'s judgments, aside from being unduly influenced by the arbitrary choice of pitch-separation, is of less practical importance than is *O*'s ability to make consistent responses to particular stimuli. If, for example, we were to regard our tones as representing nine items of knowledge about a particular continuum, we might wish to know how the pitch-separation of tones affects the amount of knowledge which *O*s receive about the continuum. Garner and Hake have shown that a measure of amount of information transmitted is useful for this purpose.⁸

Fig. 3 plots the amount of information which each mel-group receives about its stimulus-series during eight test-weeks. The left hand ordinate represents the absolute transfer of information in bits per tone, while the right hand ordinate gives the efficiency of informational transfer in percentage (*i.e.* amount of information received over amount theoretically possible with nine stimulus-alternatives). These curves show several of the same characteristics noted in Fig. 1; namely, the clear differentiation of mel-groups, failure to reach the limit theoretically possible, differences in the shape of the learning function for the wider and narrower mel-groups, and the positive relation between mel-separation and benefit from practice. Of special interest is the fact that amount of transferred information varies directly with the size of the pitch-separation of tones in the series. Varying the separation by a factor of six (*i.e.* 300/50 mels) increases the efficiency of informational transfer from 40 to 70% at the end of training.

This is not in keeping with Pollack's finding that the transfer of information "changes only a few percentages, even when the ratio between successive steps in the series is varied by a factor as great as 20 times." That this discrepancy in findings is due primarily to the extensive training given the *O*s in the present study is sug-

⁷ Mull, *op. cit.*, 479-482.

⁸ W. R. Garner and H. W. Hake, The amount of information in absolute judgments, *Psychol. Rev.*, 58, 1951, 446-459; Hake and Garner, The effect of presenting various numbers of discrete steps on scale reading accuracy, *J. Exper. Psychol.*, 42, 1951, 358-366.

⁹ Irwin Pollack, The information in elementary auditory displays, unpublished paper read before The Acoustical Society of America, May 8, 1952.

gested by the fact that the overall range of efficiency of informational transfer to our mel-groups on the initial week of training is 10%, a value which is in reasonably good agreement with the small percentage reported by Pollack.

The precise relation between interstimulus pitch-distance and information received at the end of training is shown in Fig. 4. Under the condition

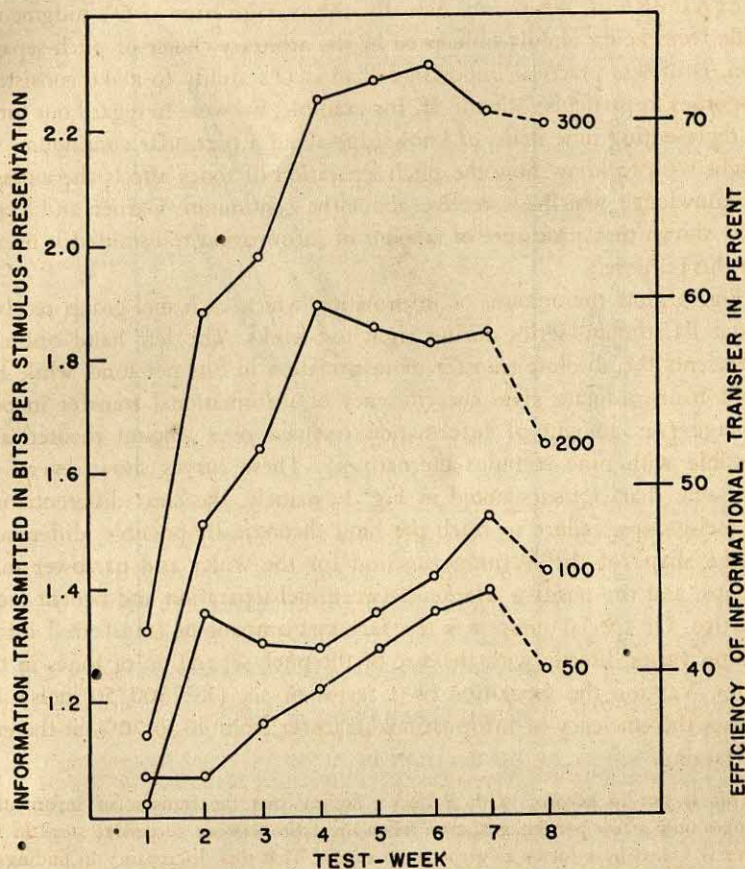


FIG. 3. AMOUNT OF INFORMATION IN ABSOLUTE JUDGMENTS OF NINE PURE TONES DURING EIGHT TEST-WEEKS.

Values for Week 5 are interpolated for reasons explained in legend of Fig. 1.

in which *O*'s judgments are corrected, the amount of information increases in an essentially linear fashion with the pitch-separation of tones. When correction is abandoned, however, the transfer of information is reduced in a way which suggests that the wider the pitch-separation, the smaller

is the informational loss. As we have already noted when considering average error, the 200 mel-group fails to conform with this tendency.

Although Fig. 4 gives every suggestion that a larger pitch-separation will result in still further informational transfer, several considerations advise against one's extrapolating this function. For one thing, above and below a frequency band extending from about 1000 to 6000 $\sim$, a given

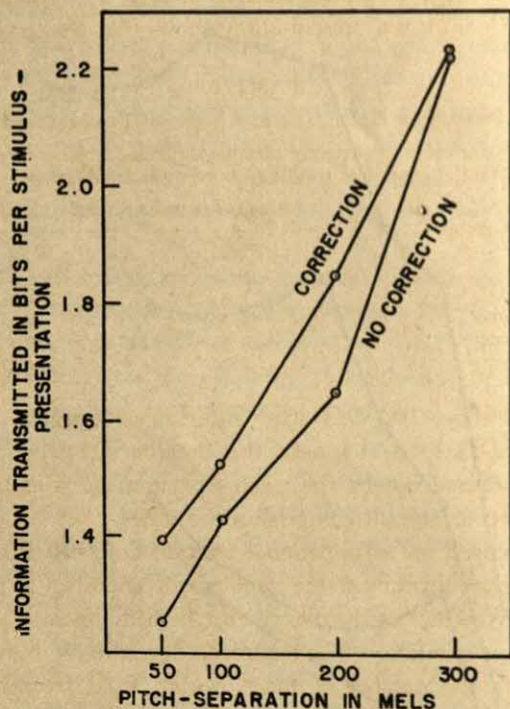


FIG. 4. AMOUNT OF INFORMATION IN ABSOLUTE PITCH-JUDGMENTS AS DETERMINED BY THE PITCH-DISTANCE BETWEEN NEIGHBORING TONES. The upper line represents the final week of corrected performance; the lower line, the final uncorrected performance.

mel-separation eats up the audible frequencies at an increasing rate. In addition to thus facing the "unhappy prospect of running out of audio-frequencies, one encounters certain practical difficulties in controlling the stimulus for pitch at the frequency extremes.

Number of tones absolutely identifiable. Within specified frequency limits, the number of tones identifiable by the method of absolute judgment will depend, in large measure, on the criterion of accuracy one

chooses. By Wedell's criterion, which consisted of a single perfect identification of the stimulus-series,¹⁰ the *Os* (three in number) in our 300-mel group had learned their series in 1, 2, and 7 weeks, respectively, and two of them could still meet this criterion when uncorrected. Two of the 3 *Os* in the 200 mel-group learned their series, by the same criterion, in four weeks, one of them being able to maintain this level uncorrected. Wedell has stated that average *Os* cannot learn to identify 13 tones about 111 *DLs* of pitch apart.¹¹ Since the present study shows that *Os* can learn to iden-

TABLE I

THE RESPONSE PROFICIENCY EQUIVALENT TO AMOUNT OF TRANSFERRED INFORMATION

Values represent the number of perfectly identified stimuli necessary for an informational transfer equivalent to that received by each mel-group.

Mel-group	Number of tones perfectly identified	
	Week 1	Week 8
50	2-3	2-3
100	2-3	2-3
200	2-3	3-4
300	2-3	4-5

tify nine tones about 106 *DLs* apart and, under certain conditions, nine tones about 70 *DLs* apart, it appears that the difficulty experienced by Wedell's *Os* was caused not by the pitch-separation of stimuli, but by the absolute number of stimuli comprising the series.

Pollack employed an informational treatment of his data on absolute pitch,¹² hoping to circumvent the problem of choosing an arbitrary criterion of achievement, but discovered that he had chosen what was essentially a zero-error criterion.¹³ The treatment, nevertheless, provides a means for determining the number of tones, within fixed frequency-limits, to which *Os* are able to respond in an unequivocal manner. Table I summarizes for the first and final weeks of testing the response-proficiency equivalent to the amount of transferred information. These values answer the question: How many perfectly identified stimuli would it take to yield an informational transfer equivalent to that received by each mel-group?

It will be seen that on the first week of testing, no mel-group receives

¹⁰ Wedell, *op. cit.*, 500. Actually the present criterion is more rigorous than Wedell's, since stimuli were so arranged that a single perfect repetition always occurred in a context of more than nine correct judgments.

¹¹ *Ibid.*, 501.

¹² Pollack, *op. cit.*

¹³ While the information measure is insensitive to the magnitude of errors (Hake and Garner, *op. cit.*, 363), it is exceptionally sensitive to the number of errors.

more information from a 9-tone series than can be gotten from a 2- or 3-tone series, perfectly identified. This remains true for the 50 and 100 mel-groups throughout the entire experiment. Eight weeks of testing, on the other hand, enables the 200 and 300 mel-groups to boost their efficiency to between 3-4 and 4-5 tones, respectively.

As might be expected, the *O*s within any one mel-group differ in regard to the amount of information each receives. The response-proficiency equivalent to the information received by the best *O* in the 300 mel-group, for example, is perfect identification among 6-7 tones. Though slightly higher than the estimate based on the combined judgments of the three *O*s, this value still places the number of absolutely identifiable tones quite low. In this respect, the present results substantiate those of Pollack, who found that the information in combined absolute judgments rarely exceeded perfect identification among more than five tones, despite changes in the total number of tones, their pitch-separation, and, to some extent, *O*'s experience in judging.¹⁴

Forgetting. How much knowledge about the stimulus-series is lost in two months during which no absolute judgments were made?

Mull reports that the learning shown by her *O*s persisted for a month.¹⁵ Wedell's *O*s, after from 2-4 months' rest, relearned their series in less than one-third the time required initially, thus indicating a substantial savings effect.¹⁶ Gough noted some practice effects persisting over a period of a year and, furthermore, that the amount of ability retained was in some measure proportional to the degree of ability acquired through practice.¹⁷ Inasmuch as we have already noted that the informational benefit from practice is positively related to the size of the pitch separation between tones, Gough's findings suggest that the larger the pitch separation, the less information will be lost following a rest.

Fig. 5 plots the amount of information received by each mel-group on the final week of testing and the amount received two months later. It will be seen that all mel-groups show an informational loss after a two-months' rest. This loss, with the exception of that for the 50-mel group, is inversely proportional to the size of the pitch-separation. Two months' rest has a practically negligible effect on the transfer of information to the 300-mel group, while the same period of rest reduces the information received by the two narrowest mel-separations to a value only slightly greater than that obtained on the first week of testing. Thus, it appears that the stabil-

¹⁴ Pollack, *op. cit.*

¹⁵ Müll, *op. cit.*, 486.

¹⁶ Wedell, *op. cit.*, 498.

¹⁷ Gough, *op. cit.*, 86.

ity of judgment, during a two-month rest-interval as measured by resistance to a change in information received, shows a tendency to increase with the size of the pitch-distance between the tones in the series.

Effect of series. A question which arises in connection with the particular spacing of stimuli in the present study is whether pairs of tones which are equidistant in pitch are about equally confused in the absolute judgmental situation. An approach to this matter is provided by a partial overlapping, among three of the four mel-series, of the particular frequencies used as stimuli. These frequencies (1330, 1700, and 2104 ~) are sepa-

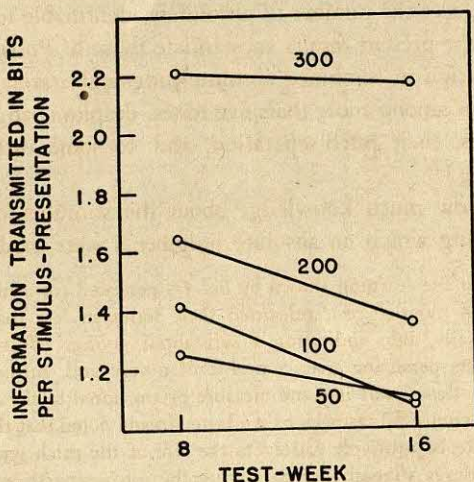


FIG. 5. EFFECTS OF A TWO MONTHS' REST ON ABSOLUTE PITCH IDENTIFICATION. Parameter is interstimulus pitch-distance in mels.

rated by 200 mels, yet occupy different positions in the 50-, 100-, and 200-mel series. For the 50-mel group, the frequencies are tones 1, 5, and 9; for the 100-mel group, tones 3, 5, and 7; and for the 200-mel group, tones 4, 5, and 6. Should all these tones be about equally confused, we could conclude that confusion is due to a given pitch-separation, irrespective of the position of the tones in the series. Should the degree of confusion vary systematically, however, with the position of the tones in the series, we would be forced to conclude that pitch-separation is only one of the factors causing confusion in a serial situation—and not the most important one. Table II, which summarizes the extent to which the three frequencies are confused on the final uncorrected week of testing, makes it clear that stimuli separated by equal pitch-distances are confused accord-

ing to their position in the judgmental series. For example, tones 1, 5, and 9 are never confused by the 50-mel group, while the same frequencies (tones 4, 5, and 6) are confused almost one-third of the time by the 200-mel group.

Other writers have reported on these so-called *anchoring effects* which characterize absolute judgments of stimuli in a series. Wedell's graphs show that the average DL-error in judging pure tones is least at the extremes of the range.¹⁸ Wever and Zener maintain that O's ability to make absolute judgments derives from a rapidly-formed

TABLE II
CONFUSION OF EQUAL PITCH-DISTANCES (200 MELS) BY THE 50-, 100-,
AND 200-MEL GROUPS

Mel-group	Stimuli	No. confusions	% confusion
50	1-5	0	0.0
	5-9	0	0.0
	1-9	0	0.0
100	3-5	7	7.0
	5-7	4	4.0
	3-7	1	1.0
200	4-5	20	21.0
	5-6	27	28.0
	4-6	3	3.0

body of knowledge about the character of the stimulus series; also, that an early acquaintance with the extreme stimuli facilitates the growth of such knowledge.¹⁹ Riker, upon observing that Os allegedly having absolute pitch performed no better at the extremes of the series than did Os allegedly without this ability, concluded that both groups of Os were about equally successful in using the ends of the range as frames of reference.²⁰ Pollack investigated the possibility that the increased accuracy noted at the extremes of his series might be due to a special identifiability factor inherent in the particular frequencies used.²¹ Partitioning the frequency range so as to exclude the centermost frequencies did not, however, appreciably increase the amount of information transmitted.

The anchoring effects noted in the present study, in view of the equal spacing of pitch-stimuli, have an important implication concerning the generality of the mel-scale for pitch. It is recognized (Garner,²² Stevens and Volkmann²³) that sensory scales may be erected by a number of different procedures; also, barring the opera-

¹⁸ Wedell, *op. cit.*, 495.

¹⁹ E. G. Wever and K. E. Zener, The method of absolute judgment in psychophysics, *Psychol. Rev.*, 35, 1928, 466-493.

²⁰ B. L. Riker, The ability to judge pitch, *J. Exper. Psychol.*, 36, 1946, 331-346.

²¹ Pollack, *op. cit.*

²² W. R. Garner, Some statistical aspects of half-loudness judgments, *J. Acoust. Soc. Amer.*, 24, 1952, 153-157.

²³ S. S. Stevens and John Volkman, The relation of pitch to frequency: A revised scale, this JOURNAL, 53, 1940, 329-353.

tion of constant errors associated with the various procedures, that the resulting scales must be in reasonable agreement to be considered valid. Now, a given degree of confusion among tones judged by the method of absolute judgment appears to constitute a legitimate criterion upon which to erect a scale of equal pitches. If this is so, then the present results indicate that the mel-scale is not a valid scale of equal pitches for series judgment operations. Similarly, Garner has shown that a scale of equal discriminability erected on data obtained from absolute judgments of loudness does not agree with the loudness function based on other than series-judgment procedures.²⁴

Cross-sensory limits. Both Miller²⁵ and Pollack²⁶ have called attention to the apparent discrepancy between the estimated number of tonal *jnds* (about 340,000, according to Stevens and Davis²⁷) and the number seemingly manageable by the method of absolute judgment. Is this apparent inability to structure the pitch-dimension along absolute lines unique, or do we encounter similar difficulty with loudness and, indeed, with unidimensional experiences in other sense modalities? Pollack believes that *Os* have difficulty keeping track of more than 5 or 6 arbitrary designations along any one continuum, although he offers no evidence to support his view.²⁸ A few such studies may be found, however. Garner reports that the amount of information *Os* received about a series of 20 different loudness steps was roughly equivalent to perfect identification among only 3-4 stimuli.²⁹ Apparently, our limitations in structuring the dimension of loudness along absolute lines are of nearly the same order as are those for pitch, being smaller if anything.

A study dealing with absolute judgments in the sphere of spatial vision is reported by Hake and Garner.³⁰ The object was to determine the minimal number of different pointer positions which can be used in a standard interpolated interval to transmit the maximal amount of information about a given event continuum. The informational transfer in this case became asymptotic with ten pointer positions, a figure somewhat larger than the five which Pollack reports for pitch.

Finally, Guilford, in discussing the optimal number of steps to be used in rating scales, notes that various investigators have placed the critical number of steps between 5 and 9, depending on a number of considerations, including the training of the rater.³¹

These findings, while fragmentary, suggest that the normal organism's capacity to structure experience in the absolute manner is crudely developed, provided we limit our consideration to experiences occurring along a single sensory dimension. From the point-of-view of biological survival, however, this inability appears not too distressing. The bulk of stimuli which individuals must unequivocally identify in an absolute manner involves the simultaneous stimulation of several experiential

²⁴ W. R. Garner, An equal discriminability scale for loudness judgments, *J. Exper. Psychol.*, 43, 1952, 232-238.

²⁵ G. A. Miller, *Language and Communication*, 1951, 49-50.

²⁶ Pollack, *op. cit.*

²⁷ S. S. Stevens and Hallowell Davis, *Hearing*, 1938, 152-154.

²⁸ Pollack, *op. cit.*

²⁹ Garner, *op. cit.*, *J. Exper. Psychol.*, 43, 1952, 238.

³⁰ Hake and Garner, *op. cit.*, 358-366.

³¹ J. P. Guilford, *Psychometric Methods*, 1936, 267-268.

dimensions.²² Requiring *O*, then, to perform certain operations, viz., absolute judgments of pure tones, may be tantamount to requiring a type of performance he rarely, if ever, undertakes. Future research, in addition to carefully exploring the organism's capacity to make absolute discriminations in every experiential dimension, must undertake the problem of determining how judgmental accuracy improves with the degree of patterning in the stimulus.

SUMMARY AND CONCLUSIONS

Four groups, each containing 3 *O*s, identified by number each of 9 pure tones 48 times a week for 8 weeks. For any group, the 9 tones were separated by equal appearing pitch-intervals (either 50, 100, 200, or 300 mels). For the first seven weeks, the correct answer was given after every judgment. No correction was given on the final week. Forgetting was measured after two months.

With relatively small pitch-separations, learning takes place slowly and is incomplete even after eight weeks; with wider pitch separations, improvement is rapid and complete within comparatively few practice sessions. Within certain limits, all three of the following are positively related to the magnitude of the pitch-distance between tones comprising the judgmental series: (1) the number of subjective reference points which *O* can set up between two limiting frequencies; (2) the rapidity with which these points can be established; and (3) their stability. It appears, however, that the number of tones which *O* can handle adequately by the method of absolute judgment is quite small, perhaps 10, or less, by a reasonable criterion. Certain indications suggest similar limitations for other modalities.

²² Miller (*op. cit.*, 52), has stressed the rôle played by absolute identifications in the perception of the complex sounds we call speech.

CONTIGUITY VS. DRIVE-REDUCTION IN CONDITIONED FEAR: THE PROXIMITY AND ABRUPTNESS OF DRIVE-REDUCTION

By O. H. MOWRER and L. N. SOLOMON,
University of Illinois

This study attempts to determine the conditions under which the response of fear becomes attached to an initially neutral stimulus. Many psychologists, following the lead of Hull,¹ currently hold that all learning depends upon the occurrence of some form of drive-reduction. Since the development of a functional connection between a formerly neutral stimulus and the response of fear is a form of learning, these psychologists, to be consistent, must hold that the development of such a connection likewise is dependent upon drive-reduction.

In contrast, others, including the present authors, believe that the drive-reduction formula, while useful in accounting for the selective learning of skeletal responses ('habits'), is not adequate to account for certain other forms of learning—conspicuous among which is that involved in the acquisition of fear. Here, it appears, the conditions of reinforcement are provided by stimulus-contiguity.²

Since one of the best ways of causing a neutral stimulus to acquire the capacity to elicit the response of fear is to pair that stimulus with an electric shock, the difference between the two theoretical positions just described can be reduced to very simple and very specific terms by saying that for Hull and his followers this type of learning depends upon what happens

* Accepted for publication October 22, 1952. The writers are indebted to Dr. Arthur L. Irion for suggestions regarding the design of the experiment.

¹ C. L. Hull, *Principles of Behavior*, 1943, Chapter 6.

² Some writers, unjustifiably in our opinion, make a distinction between 'contiguity' and 'reinforcement'; identifying the latter with drive-reduction and the former with sheer stimulus-coincidence. 'Reinforcement,' it should be remembered, literally means a strengthening and is therefore equatable to learning itself. The term is thus generic and should not, we feel, be restricted to any particular type of learning. If there are two forms of learning—that produced by drive-reduction (solution-learning) and that produced by contiguity (sign-learning)—then there are two forms of reinforcement: that produced by drive-reduction and that produced by contiguity. Historically, the justification for using 'reinforcement' to refer exclusively to drive-reduction is especially weak since the term was initially introduced by Pavlov to refer, not to the learning produced by drive-reduction, but to that developed through contiguity, or so-called conditioning. (see O. H. Mowrer, *Learning Theory and Personality Dynamics* 1950, Chapters 9 and 10; Two-factor learning theory: Summary and comment, *Psychol. Rev.*, 58, 1951, 350-354.)

when the shock *terminates*, whereas for those holding the other point of view this type of learning is dependent upon what happens at shock *onset*.

One of the earliest attempts to test differential predictions based on these two theories was an experiment carried out by Suter and Mowrer in 1947.³ In one group of Ss (laboratory rats), a signal was turned on for 5 sec., then, turned off and immediately followed by an electric shock for 10 sec. In a second group of Ss, the same type of signal (a blinking light) was turned on for 5 sec. but then, instead of being turned off, was left on throughout the 10-sec. period of shock. In the first group, the signal thus coincided with shock onset but not with shock termination, whereas in the second group, the signal coincided with onset *and* termination. It was reasoned that if shock termination were the crucial event in connecting the response of fear to an initially neutral stimulus, then that response would become more strongly connected with such a stimulus if the stimulus coincided with shock termination than if it ended at shock onset. The fact that the two groups of animals in this experiment showed no difference in the amount of fear they displayed in response to the signal in the two situations argues against the view that shock termination is a significant event in such learning. Such a finding favors the view that learning of fear, in a situation of this kind, is dependent upon the coincidence of signal and shock onset.

Several objections can be, and indeed have been, urged against this experiment, but it sharpened the issue and suggested further research.⁴

As an advance over the Suter-Mowrer experiment, a second experiment was described by Mowrer as follows:

If the reinforcement for fear learning occurs when the noxious stimulus which is forecast by a signal terminates, and if that signal lasts only until the *onset* of the noxious stimulus, then there should be poorer fear-learning when there is a long interval between the onset and end of the painful stimulus than when this interval is of short duration. In other words, under the stated conditions, a signal which is followed by a *brief* painful stimulus should acquire the capacity to elicit a *stronger* fear response than should a signal which is followed by a painful stimulus of prolonged duration. However improbable this inference may sound, it follows logically from the stated premises, for we know from other studies that in conditioning, other things equal, the briefer the interval between a stimulus and the reinforcing state of affairs, the greater the learning.⁵

There is another possibility of testing the view that fear learning is dependent upon what happens when the noxious, or 'unconditioned,' stimulus terminates. In the November, 1949, (mimeographed) version of Hull's formal system of postulates, there is a corollary of an earlier proposition which runs as follows:

When neutral stimuli are repeatedly and consistently associated with the evocation of a primary or secondary drive and this drive undergoes an *abrupt diminu-*

³ Mowrer, 1950, *op. cit.*, Chapter 10.

⁴ N. E. Miller, Comments on multiple-process conceptions of learning, *Psychol. Rev.*, 58, 1951, 375-381; Learnable drives and rewards, *Handbook of Experimental Psychology*, 1951, 446; J. J. Kendler, Reflections and confessions of a reinforcement theorist, *Psychol. Rev.*, 58, 1951, 368-374; F. D. Sheffield, The contiguity principle in learning theory, *Psychol. Rev.*, 58, 1951, 362-367.

⁵ Mowrer, 1951, *op. cit.*, 351.

tion, the hitherto neutral stimuli acquire the capacity to bring about the drive stimuli (S_D) which thereby become the condition (C_D) of a secondary drive or motivation (*italics added*).⁶

By way of commentary on this statement, Mowrer replied as follows:

What this statement means, in part, is that when a neutral stimulus is 'associated' with the occurrence of a primary drive, such as electric shock, the neutral stimulus will become capable of eliciting a secondary drive or motive—in this case, fear. This much of the statement is manifestly true. However, the implication that the learning here involved is dependent upon an 'abrupt diminution' in the primary drive is questionable. The results of the researches [of Suter, and Traum and Horton] just reported indicate that it is drive *onset*, rather than drive 'termination,' which is the important event in this kind of learning. Believing that drive diminution is the crucial event in this connection, Hull states that the acquisition of a secondary drive (such as fear) will occur only, or at least optimally, when this diminution is *abrupt*. I predict that experimentation would show that fear conditioning is not significantly influenced by the abruptness, or the lack of abruptness, with which the primary drive (unconditioned stimulus) is terminated. On the other hand, I should expect that fear conditioning *would* be influenced by abrupt vs. gradual onset of the primary drive. Confirmation of the first or, more particularly, of both of these predictions would discredit Hull's first corollary. Such confirmation would add support to the assumption that reinforcement in the sense of drive reduction is not relevant to the explanation of secondary-drive acquisition and to the assumption that there are two instead of only one type of reinforcement: effect [drive reduction] and contiguity.⁷

• The investigation which is described in detail below was therefore designed to test two inferences which flow directly from Hull's systematic position: (1) that a signal which is followed by an electric shock of brief duration will arouse a stronger fear than will a signal which is followed by a shock of long duration; and (2) that a signal which is followed by a shock that terminates abruptly will arouse a stronger fear than will a signal which is followed by a shock that terminates gradually. Neither of these inferences is supported by our findings.

METHOD AND PROCEDURE

Subjects. The Ss used in this study were 40 laboratory rats from the University's colony. They consisted of 37 albino and 3 hooded animals, ranging in age from 3 to 6 mo. Each of the four experimental groups contained five male and five female animals, with age and other variables balanced. The four groups will here be designated as Groups I, II, III, and IV.

Apparatus. The apparatus employed in this experiment has been described in detail elsewhere.⁸ Briefly, it consists of an elongated box, the floor of which is a grid which can be charged electrically. At the left end of the box is a Skinner bar and food-delivery mechanism; but for present purposes this mechanism was disconnected (to avoid noise) and the food (small pellets of Purina Laboratory Chow) was

⁶ Hull, *Essentials of Behavior*, 1951, 25.

⁷ Mowrer, *Learning Theory and Personality Dynamics*, 1950, 293-294.

⁸ Mowrer and Miller, A multi-purpose learning demonstration apparatus, *J. Exper. Psychol.*, 31, 1942, 163-170.

delivered manually into the eating trough whenever the bar was depressed sufficiently to cause a small light (visible only to the experimenter) to glow. On the interior rear wall are two $7\frac{1}{2}$ -w. electric lamps which provide illumination and which, when blinked on and off at the rate of 4~ per sec., constituted a very satisfactory stimulus for conditioning purposes.

In the present investigation, this apparatus was divided, by means of a vertical partition, into two parts, each of which contained one of the two electric lights referred to. The compartment thus formed on the right was used for associating the blinking light with electric shock (fear conditioning), while the left compartment was used for training and testing the Ss in a bar-pressing habit which will be described more fully below. To make the two halves of the apparatus strikingly different, a false interior (four walls and ceiling) of white cardboard was inserted on the right side, which contrasted with the walnut-stained wooden interior on the left. To increase further the contrast, the floor-grid in the left compartment was covered with a piece of plywood. The plate glass comprising the front of the apparatus was covered on the left side with colored cellophane, thus providing one-way vision into that half of the apparatus when illuminated from the interior. Observation of S in the right compartment was made possible by means of a mirror placed at an angle beneath the floor-grid of that compartment.

The electric shock administered from the floor-grid of the right compartment consisted of 165 v. of 60~ alternating current with a limiting resistance of 175,000 ohms, applied for the varying periods of time described in the following section. In order that the shock could be turned off gradually as well as abruptly, a potentiometer (Mallory Midgetrol U-53, with modified logarithmic winding) was put into the circuit and was so arranged that it could be driven at either of two speeds by a synchronous motor. During application of shock S was prevented from leaping off the floor-grid by means of a plexiglas 'bridge' which was placed over it.

Training procedure. All of the Ss used in this experiment were treated exactly alike except for the nature of the electric shock to which they were exposed. They were deprived of food for 48 hr. and then, individually, were put into the left side of the apparatus. Here they found a small metal bar which, when slightly depressed, prompted E to place a pellet of food in the eating trough. On the first day S was left in the apparatus until it had pushed the bar and obtained food 20 times. It was then returned to its cage and satiated by having access to wet mash for 30 min. At the end of this time it was placed back in the apparatus, but now in the right-hand compartment, where, at the end of 5 min., the light blinked for 3 sec. and was followed by shock as indicated in Fig. 1: *i.e.* Ss of Group I received a 3-sec. shock, with abrupt onset and abrupt termination; those of Group II a 10-sec. shock, with abrupt onset and abrupt termination; of Group III a 4-sec. shock, with abrupt onset and gradual termination; and of Group IV a 7-sec. shock, with abrupt onset and gradual termination.

Five minutes after the appropriate signal-shock combination was administered, S was removed from the right side of the apparatus, returned to its cage, and deprived of food for 23 hr., when it was again put back into the left side of the apparatus. Having been shocked the preceding day on the right side of the apparatus, S now had some tendency to show a generalized fear-reaction; but its hunger soon drove it to seek and obtain food by bar-pressing. On this, the second day of training, every S was allowed to obtain 5 pellets of food (as opposed to 20 on the first day), returned

to its home cage, satiated, and then given a fear-conditioning trial in the right side of the apparatus, just as on the first day. The procedure employed on the second day was repeated on each of the next three days.⁹

In brief, then, the over-all training procedure was as follows. On the first day, every *S* had 20 bar-pressing trials and one fear-conditioning trial, while on the following 4 days, every *S* had 5 bar-pressing trials and one fear-conditioning trial. In all, every *S* received 40 bar-pressing trials and 5 fear-conditioning trials. In essence, what this training did was to build up a clear discrimination on the part of *S* between the left compartment, as place-where-food-is-obtained, and the right compartment, as place-where-shock-follows-blinking-light. By the fifth day,

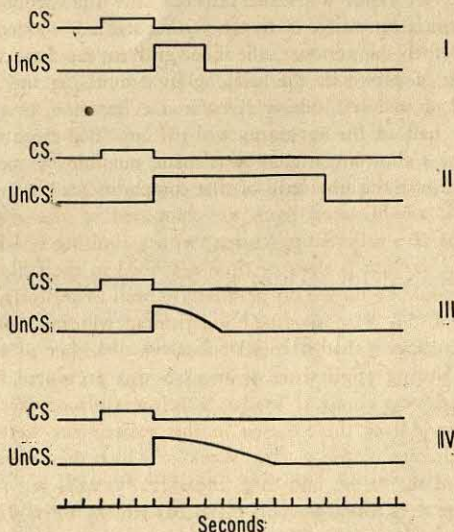


FIG. 1. SCHEMATIC REPRESENTATION OF THE DIFFERENT FORMS OF 'UNCONDITIONED STIMULUS' (SHOCK)

The signal, or 'conditioned stimulus,' was a blinking light of 3-sec. duration, which was followed in each of the four groups of *Ss* by a shock of the duration and form shown.

this discrimination was so well established that *S* pressed the bar and ate freely when put into the left compartment, whereas it was obviously frightened when put into the right compartment, and particularly so when the blinking light was presented.¹⁰

⁹ The animals were run through this training procedure in lots of four, one in each of the four groups. Ten repetitions of the total procedure were thus required to run all 40 of the animals.

¹⁰ It should be said that the animals were satiated before being put into the right compartment in order to make the experience of being in that compartment (white environment, grid-floor, no hunger) as different as possible from the experience of being in the left compartment (brown environment, wooden floor, hunger), thus facilitating the discrimination alluded to above.

Testing procedure. The object of the training procedure just described was simultaneously to set up a bar-pressing habit in response to hunger and a fear reaction to the blinking light, and then to measure the strength of the latter as a function of its capacity to inhibit the former. The test of this capacity was carried out in the following way. On the day after the 5-day training procedure, every *S* was put into the left compartment of the apparatus, with a 23-hr. hunger-drive as usual, but instead of being permitted to push the bar and obtain food only a limited number of times, *S* was allowed to push the bar and obtain food as many times as wished during an initial 5-min. period. This period of free responding established so to say, *S*'s normal response-rate.

At the end of the 5 min., the procedure was so changed that, for an additional period of 25 min., whenever *S* pressed the bar, the light which illuminated the left half of the apparatus, where *S* now was, blinked for 3 sec., just as the light in the right half of the apparatus had done on the preceding 5 days before the onset of shock. Fear acquired by the blinking light in the right half of the apparatus now generalized to the blinking light in the left half of the apparatus and acted as a punishment to *S* for having pushed the bar. The extent to which the bar-pressing response was thus inhibited was taken as an index of the extent to which the blinking light in the right half of the apparatus had acquired the capacity to produce fear.¹¹

RESULTS

The results obtained for each of the four groups of *Ss* on the test day are shown graphically in Fig. 2. As will be seen, the groups were completely indistinguishable during the 5 min. of free, unpunished bar-pressing, showing that they were homogeneous in terms of training and other characteristics. More than this, however, there is no statistically significant difference between any of the groups *after* the danger signal was introduced, thus indicating that the amount of fear conditioning occurring under the four conditions employed was roughly the same. That fear conditioning had occurred and that this particular technique afforded an efficient way of demonstrating it is indicated by the decided drop in the rate of bar-pressing as soon as the blinking light was introduced. There is no statistically trustworthy evidence that either the proximity or the abruptness of shock termination is a significant variable in this kind of learning.

¹¹ Ordinarily in laboratory research when one desires to punish behavior on the part of an organism one applies a primary drive, such as electric shock. In the present experiment what we were doing was to use a *secondary* drive; namely fear, as punishment. While constituting something of a novelty in the laboratory, this procedure is common enough in child training; parents almost universally attempt to control the behavior of their offspring, on occasion, by merely 'speaking to' or 'looking at' them. The blinking light in our experiment seems closely analogous to such parental 'warnings' or 'reminders.' Such events may appropriately be termed *secondary punishment*, just as reduction in fear is termed *secondary reward*. O. H. Mowrer, Secondary reinforcement in two-factor learning theory (in press).

As Fig. 2 indicates, the Ss of Group I were actually less inhibited by the blinking light than were the animals in the other three groups. This difference, it will be noted, is in the direction *opposite* to that demanded by Hull's formulations. The Ss of Group I received a shock which lasted only 3 sec. and ended abruptly—conditions which, if we interpret Hull correctly, would be expected to lead to maximal fear-learning! Common sense would, of course, predict the reverse; namely, that the briefer the duration of the shock the less the fear-learning, and our results are in line with this expectation. While the 3-sec. shock group does not differ

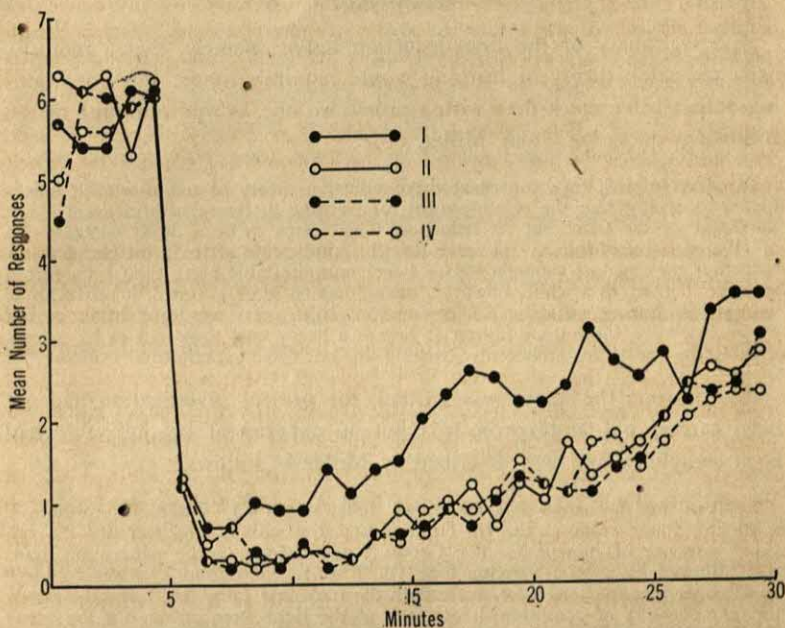
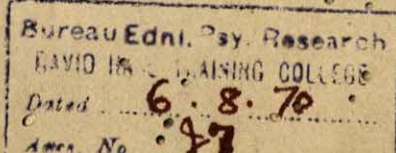


FIG. 2. CURVES SHOWING THE INHIBITION OF THE HUNGER-MOTIVATED RESPONSE. The amount of inhibition produced in Groups II, III, and IV is virtually identical, and the difference between these groups and Group I, while probably representing a real trend, is not highly significant statistically.

reliably (the 10% level of confidence) from the other three groups, it seems likely that we have here the beginning of a tendency which could almost certainly be demonstrated if the shock were made *very* brief, say, 0.5-sec. duration; it is virtually certain that it would be reliably less effective in producing fear than would a shock of longer duration.



The finding that a brief shock produces less fear-learning than does protracted shock may prove mildly troublesome for contiguity theory since in both cases the signal coincides with shock onset. The results clearly indicate, however, that fear-learning has nothing to do with what happens when the noxious stimulus of which the fear is premonitory goes off. This proposition is well established by the fact that systematic variation of both the abruptness and the proximity of termination of shock does not produce the kinds of difference in fear-learning demanded by drive-reduction theory.

DISCUSSION

The legitimacy of the first deduction tested; namely, that a monistic drive-reduction theory of learning would require stronger fear conditioning with a brief shock than with a protracted one, is well indicated by the following quotation from Miller.

In order to stick to a consistent drive-reduction theory of reinforcement, one is forced to assume that the reinforcement for learning to fear a painful event is not the onset of the pain, but its reduction. This seems to be a direct contradiction of all that is reasonable. . . . I agree that this application of the principle of drive-reduction goes against common sense; I feel uncomfortable every time I am forced to make it. Once in a while, however, when common sense is tested, it proves to be wrong: an utterly unreasonable assumption, like a ray of light being curved by gravity, or a light stone falling as fast as a heavy one, turns out to be correct. That is the reason for stubbornly insisting on conclusive experimental evidence.¹²

At the time the above was written, the present investigation had not been carried out, only projected; but an experiment by Sullivan¹³ had been completed and was described by Miller as follows.

Sullivan used a change in a pattern of lights as the CS [conditioned stimulus], a 90 db., 1000~ tone as the US [unconditioned stimulus], and recorded the galvanic response of human Ss. For Group 1 the change in the pattern of lights was followed by a brief piercing tone (0.75 sec.). For Group 2 it was followed by a longer piercing tone (4.0 sec.). Both the frequency ($p = 0.05$) and the amplitude ($p = 0.06$) of conditioned responses to the light were greater for the group having a brief, loud tone as the US. Furthermore, although there was no reliable relationship between the amplitude of the UR and the frequency of the CR, there was a relationship within each group between the rate of decrement of the UR and the frequency of the CR ($p = 0.14$ for Group 1 and 0.03 for Group 2). Thus, amazingly enough, the unreasonable deductions from a strict stimulus-reduction theory of reinforcement seem to be confirmed.¹⁴

More recently Bitterman, Reed, and Krauskopf reviewed this study.

¹² N. E. Miller, Comments on multiple-process conceptions of learning, *op. cit.* 375.

¹³ J. J. Sullivan, Some factors affecting the conditioning of the galvanic skin response, Unpublished Ph.D. dissertation, State University of Iowa, 1950.

¹⁴ Miller, *op. cit.*, 376.

with the following comment:¹⁵

Analysis of the conditioning data showed a significant difference in favor of the first group, but since (as Sullivan was careful to point out) the response level of this group was higher from the outset, the data permit no clear conclusion with respect to the validity of the hypothesis under scrutiny. In view of the initial difference between the groups, the fact that the extinction-scores showed no systematic differences may be taken as evidence against the hypothesis, but neither Sullivan nor Miller remarks upon this possibility.¹⁶

In view of these ambiguities, Bitterman, Reed, and Krauskopf have conducted a somewhat similar experiment which appears to have been better controlled and from which they conclude "that the conditioning and extinction of the GSR are unrelated to the duration of the unconditioned stimulus, at least over the range of values studied in the present experiment; the onset rather than the termination of the US seems to be the crucial factor in conditioning. This finding, which is in accord with the principle of contiguity, weighs heavily against the Hullian view."¹⁷

Like findings have been reported by Coppock and Chambers.¹⁸ In one group of 10 adult men, they recorded the extent to which the galvanic skin reaction became conditioned to a signal when the unconditioned stimulus consisted of an electric shock of 0.5-sec. duration. In another group of Ss, the same procedure was followed except that the shock lasted for 3 sec.; and in a third group it lasted for 15 sec. On the basis of their findings, these writers conclude:

We find no evidence of a gradient of reward in the conditioning of the GSR. Within the limits of the present experiment, classical conditioning occurs as well without closely associated need reduction as with it. This particular autonomic response, the GSR, can be conditioned by simply associating a neutral stimulus with the *beginning of electric shock*. This finding fails to support our interpretation of Hull's need-reduction theory of learning, but it lends support to the theories of Schlosberg, Skinner, and Mowrer.¹⁹

Presumably the way in which fear-learning is supposed to take place in terms of drive-reduction theory is as follows. The signal occurs, followed by the noxious stimulus which elicits the fear, and when the noxious

¹⁵ M. E. Bitterman, P. C. Reed, and J. Krauskopf, Conditioning and extinction of the galvanic skin response as a function of the duration of the unconditioned stimulus, this JOURNAL, 65, 1952, 257.

¹⁶ Miller does, however, suggest that the experiment should be repeated, "with a control for the possibility that the responses of Group 2 were depressed by a greater amount of negative adaptation which may have occurred during the presentations of the longer tone" (Miller, *op. cit.*, 376).

¹⁷ Bitterman, Reed, and Krauskopf, *op. cit.*, 262.

¹⁸ H. W. Coppock, and R. M. Chambers, Prior, immediate, and delayed reward in conditioning the galvanic skin response, Paper read at the Midwestern Psychological Association, Chicago, 1951.

¹⁹ *Ibid.*, 3, italics ours.

stimulus terminates it reinforces the connection between the signal and the fear as well as the connection between the noxious stimulus and the fear.²⁰ The mechanism of fear-learning is thus made essentially the same as that whereby an overt, behavioral act, mediated by the skeletal musculature, becomes more firmly connected to a drive whenever reduction in the intensity of that drive immediately follows such an act. It has been repeatedly shown, however, that reactions mediated by the autonomic nervous system (such as the galvanic skin reaction and, presumably, also fear) do not work this way. They can be conditioned to a new, initially neutral stimulus, but they cannot be learned as an instrumental, problem-solving response.²¹

The Coppock-Chambers experiment provides further evidence of this connection. Since the GSR is a long-latency response (often taking 1 to 5 sec. to reach a maximum), it is clear that an electric shock which lasts for only 0.5 sec. will be over before the response it produces has occurred. The termination of the shock cannot, therefore, reinforce the fear-response in the sense of rewarding it for having occurred. In this connection Coppock and Chambers say:

The maximal rate of decrease in skin resistance occurred about 3 sec. after each shock-onset, irrespective of shock-duration. Consequently the 1/2 sec. shock approximated rewarding the GSR by shock-termination prior to the GSR, the 3-sec. shock approximated an immediate reward, and the 15-sec. shock presented a delayed reward of shock-termination.²²

The fact that the amount of fear-learning (as represented by the magnitude of the GSR) was indistinguishable in the three groups of Ss employed in this experiment indicates, as do the results of Bitterman, Reed, and Krauskopf and our own findings, that we are here dealing with events associated with the contiguity of the signal and the onset of the noxious stimulus, and not with events associated with the termination of that stimulus. This point of view is still further substantiated by experimental findings which will be reported shortly by Mowrer and Aiken.²³

SUMMARY

In his attempt to derive a maximally parsimonious theory of learning, Hull posited that in the final analysis all learning is dependent upon the

²⁰ Hull, *Principles of Behavior*, 76 ff.

²¹ Mowrer, *Learning Theory and Personality Dynamics*, Chapter 9.

²² Coppock and Chambers, *op. cit.*, 3.

²³ Mowrer and E. G. Aiken, Contiguity vs. drive reduction in fear conditioning: Temporal variations in conditioned and unconditioned stimulus, this JOURNAL, 61, 1954, 26-38.

'primitive reinforcement' which occurs when a drive is significantly reduced or terminated. While there is good evidence that that form of learning commonly known as trial-and-error or problem-solving does indeed depend upon drive-reduction, there appears to be another, very different form of learning, which is dependent upon *drive-induction*. The purest example of the latter kind of learning is that which occurs when the response of fear becomes connected to an initially neutral stimulus as a result of that stimulus having been associated with the onset of a painful, noxious stimulus of some sort.

It is formally stated as part of Hull's system that learning occurs best (a) when the reinforcing state of affairs provided by drive reduction is temporally proximate rather than remote and (b) when reduction of the drive is abrupt rather than gradual. In the investigation here reported (40 laboratory rats were the Ss), it has been found that acquisition by an initially neutral stimulus of the capacity to elicit fear is not dependent upon either the proximity or the abruptness of drive-reduction, but is instead contingent upon what happens when the noxious, painful stimulus first occurs. This finding, supported by the results of other experiments which are cited, suggests that there are *two distinct kinds of learning*, solution-learning and sign-learning, rather than just one kind, as Hull's system presupposes.

CONTIGUITY VS. DRIVE-REDUCTION IN CONDITIONED FEAR: TEMPORAL VARIATIONS IN CONDITIONED AND UNCONDITIONED STIMULUS

By O. H. MOWRER and E. G. AIKEN
University of Illinois

The writers, who hold that all learning is dependent upon what happens when a drive is reduced, are required, by the nature of their theory to assume that when a neutral stimulus precedes or accompanies a noxious one, the neutral stimulus acquires the capacity to elicit fear because of the state of affairs provided by the *termination* of the noxious stimulus. From this it follows that, other things being equal, the closer a neutral stimulus, or signal, is in time to the termination of the noxious stimulus the more potent will it become with respect to its capacity to produce the fear-reaction. The present writers, on the other hand, believe that fear-learning is dependent, not upon what happens when the noxious stimulus terminates, but rather upon events connected with its onset.

In a preceding study these two alternative theories were tested by holding the relationship between the onset of the neutral and the noxious stimulus constant and by varying the proximity or abruptness of the termination of the latter.² In the present experiment essentially the same theoretical issue is attacked but by a different method.

As indicated in the preceding study, the question of precisely how fears are learned is of major importance in determining whether there are two basic kinds of learning or only one. If the results of the present investigation indicate that fears are learned because of what happens when a noxious stimulus occurs rather than when it disappears, there will be further evidence, in addition to that already cited, for believing that a monistic conception of learning of the type advocated by Hull is inadequate. The present experiment is also designed to bear upon Hull's corollary regarding the nature of *secondary reinforcement*.³

* Accepted⁵ for publication October 22, 1952. The writers are indebted to Dr. Arthur L. Irion for suggestions regarding experimental design.

²See C. L. Hull, *Principles of Behavior*, 1943, 422.

³O. H. Mowrer and L. N. Solomon, Contiguity vs. drive reduction in fear conditioning: The proximity and abruptness of drive reduction, this JOURNAL, 67, 1954, 15-25.

⁵Hull, Behavior postulates and corollaries, *Psychol. Rev.*, 57, 1950, 173-180.

The purpose, then, of this study is two-fold: (1) to determine whether fear-learning occurs more readily when the initially neutral stimulus to which this reaction becomes attached is temporally contiguous with the onset of shock or temporally contiguous, as Hull's formulations would seem to require, with the termination of shock; and (b) incidentally, to clarify the relationship between the conditions under which a secondary drive such as fear is acquired and the conditions under which an initially neutral stimulus becomes capable of producing, not secondary motivation, but secondary reinforcement.

METHOD AND PROCEDURE

Subjects. The Ss used in this study were 50 laboratory rats from the University's colony. They consisted of 46 albino and 4 hooded animals, ranging in age from 3 to 6 mo. Each of the five experimental groups contained five male and five female animals, with age and other variables equated. The five groups will be designated as Groups I, II, III, IV, and V.

Apparatus. The apparatus employed in this experiment is exactly the same as described in its companion study.⁴

Training procedure. The training procedure employed in this experiment was the same for all Ss except for the difference in the type of fear-conditioning they received.

The procedure on the first day for all animals was as follows. Having been deprived of food for 48 hrs. S was put into the left half of the apparatus and left there until it had pushed the bar and obtained food (a small pellet of Purina Laboratory Chow) 20 times. It was then returned to its home cage, allowed access to wet mash for 30 min., and then taken back again to the apparatus and put into the right-hand compartment. Here, after an interval of 5 min., the $7\frac{1}{2}$ -w. light which illuminated the interior of that compartment was turned on and off at the rate of four cycles per sec. for a period of 3 sec., in association with an electric shock (165-v. AC with a limiting resistance of 175,000 ohms) according to the various combinations of light and shock shown in Fig. 1 for the 5 groups. As will be seen from this diagram, Ss of Group I received the blinking light just before the shock came on; those of Group II just after the shock came on; of Group III just before the shock went off; and of Group IV just after the shock went off. The Ss of Group V served as a kind of control in that the blinking light was not temporally associated with the shock at all; instead the light occurred 2 min. after the termination of the shock. In all cases the Ss were left in the right compartment for 5 min. after the shock was terminated, making a total of 10 min. in this compartment, with shock in the middle. The Ss were then returned to their home cages and left there without food until the next experimental session, approximately 23 hr. later.

On the second day of training and the three following days, the procedure was exactly the same as on the first day, except that instead of being allowed to push

⁴ Mowrer and Solomon, *op. cit.*, 17-19.

the bar and get a pellet of food in the left compartment 20 times, each *S* was allowed to do so only 5 times. In the course of the five days of training, each thus made 40 reinforced bar-pressing responses and had five fear-conditioning trials, the latter varying as regards the particular way in which the conditioning stimulus (blinking light) was associated with the unconditioned stimulus (shock) in the five different groups, as already indicated.

The purpose of this training procedure was to set up a clear discrimination between the left half of the apparatus, as place-where-bar-pressing-produces-food

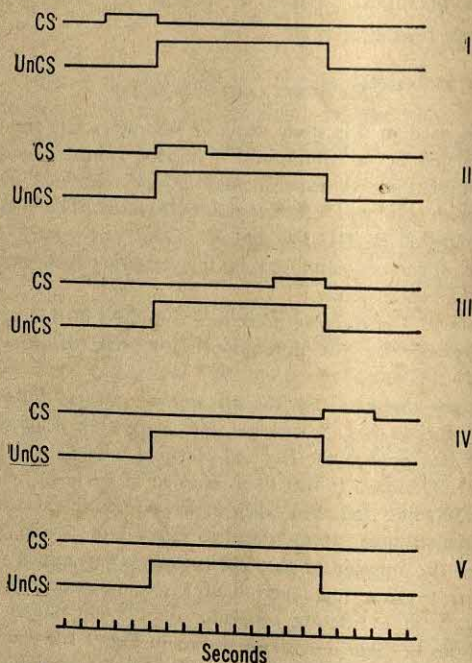


FIG. 1. SCHEMATIC REPRESENTATION OF THE FIVE DIFFERENT ARRANGEMENTS OF CONDITIONED (CS) AND UNCONDITIONED STIMULUS (UnCS)
The 'conditioned stimulus' (a blinking light) in Group V was not immediately associated with the unconditioned stimulus (electric shock) in any way, but was presented alone, 2 min. after the UnCS had terminated.

and the right half of the apparatus, as place-where-blinking-light-and-shock-occurs. The way in which this training was turned to experimental account is described below. It should be added that the *Ss* in Groups I, II, III, and IV were run in groups of four, one *S* from each group. Ten repetitions of this procedure were thus involved in the total experiment. The *Ss* in Group V were run later, as an afterthought, in two groups of five each.

Testing procedure. The central objective of this experiment was to determine what differences, if any, emerge in the conditioning of the fear-reaction as a result of having the conditioned stimulus associated with the unconditioned stimulus in the

ways specified. At the same time that this type of training was occurring, our *Ss* were also learning to press a small metal bar as a means of obtaining food. In the final, the testing part of the experiment, the fear-learning was measured as a function of its capacity to act as an inhibitor of bar-pressing. This was done in the following way.

After the five days of training, every *S*, motivated as usual by a 23-hr. hunger-drive, was put into the left side of the apparatus and allowed to press the bar and obtain food unobstructedly for 5 min. This procedure established *S*'s normal rate of response. Then, for an additional period of 25 min. every *S*, upon pressing the bar, experienced the blinking light for 3 sec. To the extent that fear had become

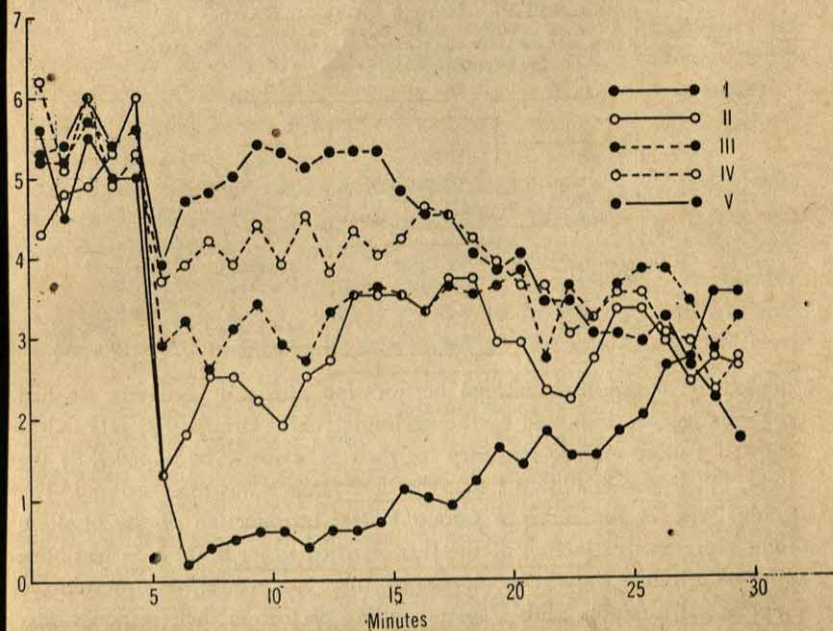


FIG. 2. CURVES SHOWING THE INHIBITION OF THE HUNGER-MOTIVATED RESPONSE

attached to the blinking light in the right compartment it generalized to the blinking light in the left compartment and acted as a 'secondary punishment' of the bar-pressing behavior.⁵ Thus the amount of depression in bar-pressing provides an index of the amount of fear-learning that has occurred during the preceding training days. The results for the five groups of *Ss* used in this experiment are shown in Fig. 2.

RESULTS

Inspection, as well as an *F*-test, indicates that our five groups of *Ss* did not differ significantly in their bar-pressing performance during the first 5

⁵ *Ibid.*, 20.

min. of the testing-period. When, however, the blinking light was introduced the groups diverged markedly. As will be seen, the Ss of Group I inhibited most strikingly, with Groups II, III, IV, and V following that order. As might have been expected, the Ss of Group V—for whom the blinking light had not been associated with the shock at all, only in the situation in which the shock occurred—showed only a slight, transitory 'dip' in their performance when the light was introduced. The

TABLE I

STATISTICAL COMPARISONS OF THE MAGNITUDE OF THE DIFFERENCES AMONG GROUPS I-V IN TERMS OF THE *T*- AND *P*-VALUES.

The raw data for these comparisons was taken from the sixth to the twentieth minute on the test-day

Groups	I	II	III	IV
II	$t=2.316$ $P=.047$			
III	$t=4.409$ $P=\text{below } .01$	$t=.838$ $P=.414$		
IV	$t=9.835$ $P=\text{below } .01$	$t=2.287$ $P=.049$	$t=1.673$ $P=.115$	
V	$t=17.414$ $P=\text{below } .01$	$t=3.636$ $P=\text{below } .01$	$t=3.484$ $P=\text{below } .01$	$t=2.699$ $P=.016$

quickly recovered their normal bar-pressing rate and gradually satisfied their hunger, as indicated by the declining curve. Groups IV, III, and II showed a more marked tendency for their behavior to be impeded by the light, but it was in Group I that this interference was most striking. The gradual rise in the curve of Group I after introduction of the blinking light represents extinction of the fear attached to it; if these Ss had been allowed a longer time in the apparatus, they, too, would have presumably approached satiation, with a corresponding decline in their response-rate.

The reliability of the differences between the various groups can be estimated in two ways. First we may note that there are 120 different ways in which these 5 groups might have arranged themselves, if chance alone were operating. The fact that the obtained order was the predicted one would thus represent a 'lucky hit' of the order of 1 in 120. Such odds are commonly accepted as indicating that factors other than 'chance' determined the outcome. Then, as a further measure of the reliability of the results obtained, we can apply the *t*-test to the difference between the mean bar-pressing scores of various pairs of groups of Ss during the crucial phase of the test-period (6 to 20 min.). Five groups of Ss make possible 10 paired comparisons, and Table I gives the results.

DISCUSSION

The results of this study show decisively that the pairing of an initially neutral stimulus with the *onset* of a noxious stimulus, such as electric shock, leads to more effective fear-conditioning than does the association of such a stimulus with the termination of the noxious stimulus. Such a finding is consistent with common-sense observations and what one would suppose to be sound biological principles, but it is inconsistent with the deductions which follow from Hull's postulates. Biologically, a 'signal' which occurs after an event has already happened is not much use, and it is therefore reassuring to find that nature's wisdom (evolution) has provided that an incidental stimulus which immediately precedes an event (in this case a noxious one)—and is thus most likely to make adaptive behavior possible—leads to more efficient conditioning than does a 'signal' which occurs later.

A 'signal' which occurs long after a noxious event is, of course, almost completely worthless to an organism; and we find, appropriately enough, that the Ss of Group V—for whom the blinking light followed the shock by 2 min.—showed only passing interest in this stimulus. Presumably such inhibition as it did produce was due to second-order conditioning. Since the 'situation' provided by the right half of the apparatus had been paired with shock, it took on some fear-producing properties; and since the blinking light occurred there, it would likewise be expected to acquire, in appropriately attenuated form, some capacity to elicit fear. Biologically, this makes sense. It is worth something to be able to be at least a little afraid of a stimulus that 'reminds' us of a situation in which we have previously had trouble. It also seems appropriate that an organism should not be very markedly disturbed by such a stimulus when it occurs in another situation (like the left side of our apparatus), in which the organism has reason to suppose it is really pretty safe.⁶

The results reported in this study make it apparent that the best place for a stimulus to occur if it is going to acquire the capacity to produce fear is just before the onset of the so-called unconditioned stimulus. They also show, however, that *some* learning will occur even though the incidental stimulus occurs *after* the onset of the noxious stimulus. Since two-factor learning theory stresses, in connection with fear-acquisition, the im-

⁶ The blinking light, while in no way painful, would undoubtedly have some distracting properties if introduced *de novo* in the bar-pressing test, but any effect of novelty was presumably eliminated by virtue of this stimulus having been experienced a number of times in the right side of the apparatus. Such inhibitory effect as it actually had was therefore probably due to second-order conditioning of the kind described above.

portance of what happens at the *onset* of shock (or other painful experience) rather than what happens at its termination, the fact just cited might be said to constitute something of an embarrassment for such a theory. It may indeed be that some qualification will eventually have to be made in respect to this emphasis upon the coincidence of CS with the *onset* of the *UnCS*; but other possibilities have not yet been thoroughly explored. One might, for example, say that while a signal gets conditioned to fear best if it immediately precedes the onset of pain, such an incidental stimulus becomes somewhat conditioned to this reaction by virtue of the pain *continuing* coincidental with the stimulus.

There is the fact that the Ss of Group IV showed some fear-conditioning despite the fact that in their case the 'signal' appeared just after the shock terminated. To be sure, this form of 'conditioning' is not very effective, but it nevertheless occurs. How can it be explained? One possibility is that it is not so much contiguity of signal and pain that is important in fear conditioning as it is contiguity of signal and the fear-reaction itself. Since the fear-reaction continues to occur rather strongly for a time immediately after cessation of the shock, conditioning might continue to be a real, though diminishing, possibility. Or one might argue that in the latter case, and perhaps even in those cases where the incidental stimulus overlaps (but does not precede) the shock (Groups II and III), the totality of environmental stimuli *i.e.* the 'situation' becomes a kind of substitute (second-order) 'unconditioned stimulus' with which the specific 'conditioned stimulus' is associated. We have already seen a slight tendency for the pairing of the blinking light and the shock-situation (right side of apparatus) to lead to some fear-learning (Group V). How much more potent the purely situational factors might be expected to be during or just after the shock!

These, obviously, are complicated questions on which we at present have very little information. In conceding the existence of this problem, we do not, however, compromise our position that it is not shock-termination, or 'relief,' which is important in fear-conditioning. Regardless of whether it is precisely the 'onset' of pain or something else that provides the reinforcement involved in fear-learning, there is no indication that pain-reduction is the crucial determinant in this connection.

Alternative interpretations. Apologists for Hull's hypothesis that all learning is ultimately dependent upon 'primitive reinforcement' (drive-reduction) have shown considerable resourcefulness in re-interpreting experimental findings which, in plain fact, go counter to this hypothesis. These re-interpretations are often put forward in highly specialized

language, but it commonly turns out that when they are stripped of their semantic complexity they have a patently *ad hoc* quality and are often mutually incompatible. We may anticipate some of the alternative proposals which will probably be made with respect to the present findings.

Miller and Dollard have long been aware of the possibility that Hull's monistic conception of learning theory may be in some respects inadequate;⁷ and in considering this possibility they have made the following suggestion, which they repeat in their more recent volume.⁸ They say:

Electrical recording of the responses of afferent nerves indicates that the sudden onset of a new stimulus produces at first a strong burst of impulses from the sense organ which rapidly diminishes in strength till a plateau of stimulation is reached.⁹ This diminution is called adaptation. According to the principles which have been outlined, such a reduction in strength of stimulation should, if marked enough (as might be the case following the sudden onset of a relatively strong stimulus), act as a reward to any responses associated with it. Careful experiments may reveal that such a mechanism accounts for certain cases of learning which might superficially appear to be exceptions to a rigorous drive-reduction theory of reward.¹⁰

Applied to the present experiment, what does this conjecture imply? It implies that reward through 'drive-reduction' may occur immediately after the onset of an electric shock, as well as at its termination. If this be true, then much of our reasoning may be impugned. There is, however, this saving circumstance: Miller and Dollard's learning theory, like that of Hull, is in essence a drive-response-reward theory; and in the present investigation we are dealing with the response of fear. This response, we know, has a relatively long latency, often as much as 1-2 sec.; whereas the Adrian 'on-effect' is measured in terms of milliseconds. Any reward which might conceivably be connected with this phenomenon would, therefore, have been over long before the fear-reaction took place. While the evidence is not yet definitive, it favors the view that responses mediated by the central nervous system (the behavioral, 'voluntary' ones) are reinforced only by rewards that follow them, *not* by rewards that precede them;¹¹ and the indications are that responses mediated by the autonomic nervous system (emotional, 'involuntary' ones) are not influenced by reward at all.¹² Thus the suggestion that acquisition by the blinking light of the ability to elicit the fear-reaction in the present investigation was dependent upon a form of reward made possible by the Adrian 'on-effect' would seem doubly doubtful: doubtful because this effect occurs *before* rather than after the fear-reaction, and doubtful because the fear-reaction seems to differ fundamentally from overt, behavioral responses in not being reinforceable by reward.

Or an argument might be advanced somewhat as follows. It might be said, in common-sense terms, that once the electrical shock came on, our rats were so

⁷ N. E. Miller and J. Dollard, *Social Learning and Limitation*, 1941, 35, footnote.

⁸ J. Dollard and N. E. Miller, *Personality and Psychotherapy*, 1950, 46, 47, footnote.

⁹ E. D. Adrian, *The Basis of Sensation*, 1928, 67, 116.

¹⁰ Miller and Dollard, 1941, *op. cit.*, 35.

¹¹ M. O. Nagaty, The effect of reinforcement on closely following S-R connections: I. The effect of a backward conditioning procedure on the extinction of conditioned avoidance, *J. Exper. Psychol.*, 42, 1951, 239-246.

¹² Mowrer, 1950, *Learning Theory and Personality Dynamics*, Chapter 9.

distracted by it that they did not notice nor attend to the blinking light.¹² If true, this conjecture would account for the good conditioning of the Ss of Group I and the poor conditioning of the Ss of Groups II and III. The Ss of Group I experienced the light *before* the shock and might, therefore, properly attend to it; but those of Groups II and III got the light after the onset of shock and might have remained fairly oblivious to it. Thus, when the blinking light subsequently occurred alone it might have been relatively meaningless.

One cannot deny to this interpretation a certain plausibility. There are however, two objections to it. First, it has some of the *ad hoc* quality previously referred to and does not, for example, explain the findings of Mowrer and Solomon, which also indicate, but in a different way, the lack of any dependence as far as fear learning is concerned upon what happens when the noxious stimulus terminates.¹⁴ Secondly, the difference in conditioning shown by Groups II and III, while small and statistically non-significant, is actually in the *reverse* direction to what the hypothesis of distraction ('neutral-interaction') would seem to imply. If fear-conditioning was really dependent upon what happens when the noxious stimulus was present, then the nearer the incidental stimulus was to the termination of the noxious stimulus, the better would fear be conditioned. Our results suggest, on the contrary, that as one moves the CS toward UnCS termination, the *poorer* the fear-conditioning becomes.

Another, perhaps rather trivial, objection that may be urged against our interpretation of the results of the present investigation is the following. Since, in the procedure of Group I, the blinking light goes off (stops blinking) just as the shock comes on, and since, in Group II, it goes on (starts blinking) just as the shock goes on, the blinking light may be said to have been 'present' at the onset of shock in both instances. Since coincidence of the CS with the onset of the UnCS is here supposed to be the crucial consideration in fear-learning, it would perhaps seem that the procedures of Groups I and II should give roughly comparable rather than strikingly different results. If we assume, however, as we reasonably may, that a 'stimulus' is a function not only of the immediate presence of its objective source but also of some accumulation of 'stimulus-traces' (in Hull's sense of the term) from prior stimulation, then we can say that the conditioned stimulus in the procedure of Group II was not 'present' at the onset of shock at all in the way it was when the conditioned stimulus came on and remained on for 3 sec. before the onset of the shock, as was the case in the procedure of Group I.

Secondary reinforcement. As indicated in the introductory section of this paper, Hull has adduced virtually the same postulates to account for the acquisition by an initially neutral stimulus of the capacity to elicit secondary motivation, *e.g.* fear, and, for the acquisition by such a stimulus of the capacity to produce 'secondary reinforcement.' For Hull this latter term

¹² More formally, the same idea could be stated in terms of Hull's hypothesis of afferent neural interactions, which runs as follows: "A given afferent receptor discharge (s_1) is modified by interaction with other receptor discharges (s_2 or s_3) on their way to or actually entering the central nervous system at about the same time," (*Essentials of behavior*, 1951, 43).

¹⁴ Mowrer and Solomon, *op. cit.*, 22.

seems to have meant—although his writings on this score are ambiguous—a conditioned form of 'primitive reinforcement'; whereas for us this term means merely the reinforcement that occurs when a secondary drive is reduced. In any case, secondary motivation and secondary reinforcement are manifestly different—not to say antithetical—and it would seem most anomalous if they developed under identical circumstances. This issue is being considered at greater length elsewhere;¹⁵ but it is appropriate here to ask what bearing, if any, the findings of the present experiment have upon it.

If we take an uncomplicated, unqualified view of the hypothesis that all learning is dependent upon drive-reduction, then we might expect that the procedure of Group III would have been the optimal one, not only for the acquisition by the CS of the capacity to elicit fear (secondary motivation), but also for acquisition of the capacity to provide secondary reinforcement. In this procedure a stimulus is "repeatedly and consistently associated with the evocation of a primary . . . drive and this drive-stimulus undergoes an abrupt diminution,"¹⁶ and the incidental stimulus is maximally near the drive diminution, thus providing, ostensibly, the best of all conditions for the acquisition by "the hitherto neutral stimulus [of] the capacity to bring about . . . a secondary drive or motivation."¹⁷ Thus, the blinking light in the procedure of Group III should have led to maximal conditioning of fear. Likewise, in this same procedure, "a neutral receptor impulse . . . occurs repeatedly and consistently in close conjunction with a reinforcing state of affairs [drive reduction]" and thus would "itself acquire the power of acting as a [secondary] reinforcing agent."¹⁸ Since the blinking light in the procedure of Group III immediately precedes the 'reinforcing state of affairs' in the sense of shock-termination, then the blinking light should *also* acquire a tendency to act as a secondary reinforcing agent.

It therefore appears that the procedure of Group III should, in Hull's system, provide the conditions under which the blinking light would have the greatest tendency to produce the secondary motive of fear and also the greatest tendency to produce secondary reinforcement. As pointed out, however, secondary motivation (drive-increase) and secondary reinforcement (drive-reduction, according to our view) are antithetical processes; and it seems most improbable that living organisms have evolved nervous

¹⁵ Mowrer, Secondary reinforcement in two-factor learning theory (in press).

¹⁶ Hull, 1951, *op. cit.*, 21.

¹⁷ *Loc. cit.*

¹⁸ *Ibid.*, 28.

systems that really work in this way. One would suppose, rather, that antithetical processes would be associated with antithetical, or at least different, conditions; and such an expectation is met if we posit a two-factor conception of learning. According to this view, the secondary motive of fear is learned through coincidence of signal with onset of pain, whereas secondary-reinforcement capacity is acquired by an incidental stimulus as a result of its coincidence with the *decrement* (drive) of pain.

This analysis does not, however, specify the nature of 'secondary reinforcement.' Hull has postulated that it involves a conditioned form of 'primitive reinforcement,' whereas we believe it is probably nothing more than what occurs whenever a secondary drive undergoes significant diminution. The question now facing us is whether the present experiment provides any evidence in favor of one as opposed to the other of these two hypotheses. We believe that it does.

If we take the position that an incidental stimulus, such as the blinking light in the present investigation, acquires maximal capacity for producing the secondary drive of fear by immediately preceding primary drive onset, then the procedure of Group I should lead to best fear-learning—and this inference is emphatically confirmed by our results. Let us now make the assumption with Hull that secondary reinforcement is a conditioned form of primary reinforcement which is acquired by an incidental stimulus which immediately precedes such a state of affairs. On the basis of this assumption one would expect the blinking light to acquire maximal properties of secondary-reinforcement in the procedure of Group III, with perhaps somewhat less good learning taking place in the procedure of Group IV. If, in terms of Hull's theory, the blinking light thus acquires properties of secondary reinforcement, it should, when presented immediately after the bar-pressing in the testing situation, *facilitate* such behavior rather than inhibit it. The results for the Ss of Groups III and IV show no facilitative influence whatever; instead, as Fig. 2 indicates, the blinking light had an inhibiting effect which, though less marked than that observed in the Ss of Group I, was still unmistakable.

What do these findings mean? Do they mean that there is no such thing as 'secondary reinforcement'? They mean, or at least they suggest, that there is no such thing as 'secondary reinforcement' in Hull's sense of the term. These results do not, on the other hand, militate against the view that secondary reinforcement is a real phenomenon if by the term we mean simply the reinforcement which occurs when a secondary drive, such as fear, is reduced. Under the conditions of the present experiment, we would hardly

have expected the blinking light to have secondary reinforcing properties during the test-period in any of our groups. All the Ss, it will be recalled, were tested under conditions in which fear was presumably minimal. In the left side of the apparatus nothing of an untoward nature had ever occurred, and the pleasure of eating had repeatedly been experienced there. We can hardly suppose, therefore, that when, at the beginning of the sixth minute on the test-day, an S pushed the bar for food and produced the blinking light that that stimulus would have reinforced bar-pressing through the fear-reduction it provided. If there is no fear to reduce, a fear-reduction (secondary reinforcement) simply cannot, according to our view, take place! Indeed, in our experiment, the blinking light had a tendency, even in the Ss of Groups III and IV, to elicit some fear, because it 'reminded' them, at least passingly, of the situation where they had been shocked, and thus produced the moderate amount of response-inhibition noted.

The foregoing discussion may seem more speculative than empirical, and we cannot deny that there is as yet comparatively little factual confirmation. Direct evidence is not, however, entirely lacking, perhaps the best of which is that recently obtained by Barlow.¹⁹ Full details on this investigation are not at present available; but a preliminary report indicates, first of all, that in carrying out an experiment something like the one here reported, Barlow has, in one important respect, obtained the same results: he has found that an incidental stimulus associated with the onset of shock is decidedly superior to one associated with the termination of shock as far as fear-conditioning is concerned. Barlow, however, has also found that a stimulus previously associated with the termination of shock is somewhat facilitative, thereby betokening the operation of secondary reinforcement.

The question thus arises: Why does Barlow find that an incidental stimulus associated with shock-reduction acquires secondary reinforcing properties, whereas in our experiment no such effect was obtained? The answer, it seems, lies in the circumstances under which the effects of such a stimulus were tested. As already indicated, the present investigators tested the effects of the incidental stimulus under conditions in which the Ss were presumably not already afraid and thus were not capable of experiencing secondary reinforcement in the sense of fear-reduction. Barlow, on the other hand, tested his Ss under circumstances in which they had reason to be somewhat apprehensive, with the result that a stimulus which had previously been associated with shock-termination was interpreted by his Ss as meaning that now 'everything is going to be all right again.' In other words, in Barlow's experiment, the stimulus which had been associated with shock-reduction was evidently capable, under test-conditions, of providing a measure of reassurance to the Ss, whereas, in our investigation, a stimulus with a similar history did not have such an effect because the test-conditions were different.

¹⁹ J. A. Barlow, Secondary motivation through classical conditioning: One-trial non-motor learning in the white rat, *Amer. Psychol.*, 7, 1952, 273.

This finding must be interpreted provisionally until fuller data are available; but it is consistent with other studies in suggesting that Hull's conception of secondary reinforcement, as something which can occur independently of secondary-drive reduction, is debatable.²⁰ It favors moreover, our view that secondary reinforcement is, purely and simply, the reinforcement that occurs when a secondary drive is reduced or abolished.

SUMMARY

As indicated earlier, the purpose of the experiment here described was two-fold: (a) to determine whether learning occurs more readily when an initially neutral stimulus is temporally contiguous with the onset of shock or is temporally contiguous, as Hull's formulations would seem to require, with the termination of shock; and (b) incidentally, to clarify the relationship between the conditions under which a secondary drive, such as fear, becomes attached to a neutral stimulus and the conditions under which such a stimulus becomes capable of producing, not secondary motivation, but 'secondary reinforcement.'

The results obtained confirm the view that fear-learning occurs more readily (only?) when the initially neutral stimulus is contiguous with the onset of shock than when it is contiguous with the termination of shock. The results also indicate, though less directly and less definitely, that secondary reinforcement consists merely of secondary-drive reduction and not, as many of Hull's statements imply, of a conditioned form of 'primitive reinforcement,' which purportedly can operate without drive-reduction.

These findings, like those of the preceding study by Mowrer and Solomon, are consistent with the view that there are two basically different forms of learning: sign-learning (conditioning, stimulus-substitution, place-learning) and solution-learning (trial-and-error, response-substitution, problem-solving, act-learning). No monistic conception of learning, including Hull's, seems capable of accounting for the presently known facts in a reasonably inclusive and consistent manner.

²⁰ Mowrer, Secondary reinforcement in two-factor learning theory (in press).

A TEST OF THE HYDRAULIC THEORY OF COCHLEAR MECHANICS THROUGH MULTIPLE FREQUENCY BEATS

By MAX F. MEYER, Miami, Florida

The present article reports what musical observers testify they hear when they listen to tones of various ratios and it relates their testimony to auditory theory—in particular to the hydraulic theory. This article is a complement to my recent study on cochlear mechanics, acquaintanceship with which is presupposed.¹

Any theory of cochlear mechanics must explain facts of psychophysiological experiments. In the acoustic literature extra psychological (albeit biological) microphonics (derivable from the head of a cat—or the head of a cabbage when vehemently shaken) are found substituted for painstakingly listening. Such records of microphonics are scarcely a contribution to the theory of hearing. Hearing equals listening.

Vibratory (place) theories quite commonly abstain from asking the student to listen for himself. Instead, *the student is told what he hears*.²

In the previous article I pointed out why, for reasons of pure mathematics, a doubling of the beats of the higher tone in a mistuned multiple

* Accepted for publication May 24, 1952.

¹ M. F. Meyer, Crucial experiments in cochlear mechanics, this JOURNAL, 66, 1953, 261-269.

² For example, S. S. Stevens and Hallowell Davis, *Hearing: Its Psychology and Physiology*, 1938, 1-489, write: "Tell him [the student] what he hears . . . then he is in a position to be told, beginning with Chapter 10, how he hears it" (p. ix); "When the ear is stimulated by a pure tone, we [the reader is told] hear . . . aural harmonics. . . . When two loud tones are sounded together, we hear . . . the sums and differences" (p. 184): [The book mentions no one by name as testifying that he actually hears these aural harmonics, sums, and differences.] "The microphonic action . . . may be detected from almost any part of the head when sufficient amplification is used" (p. 312). "A small section of normal organ of Corti . . . will generate some potential at all frequencies" (p. 338). [That this contradicts *pitch* placement in the organ of Corti is not acknowledged.] "The piezo-potentials of the hair-cells . . . are possibly . . . an epiphenomenon like the noise of an automobile" (p. 346). [How then do they explain hearing?] "Of this we may be certain: the position of basilar disturbance is related to the perceived pitch" (p. 375). [This is said to have been demonstrated by Reboul.] "Reboul was able to show . . . maximal displacement at a position which is a function of frequency" (p. 444). [This is a misquotation, Reboul actually found a lack of localization. See this JOURNAL, 63, 1950, 276, for the correct quotation from Reboul. The above samples from the book show that its chief interest lies in recording biological microphonics. In spite of its title it is devoid of reports of what any definite listeners consciously 'hear' when exposed to a combination of physical tones.]

illustrating the hydraulic theory) of the maxima and minima of the following two equations. Phase III gives the same results as Phase I; and Phase IV the same as Phase II. Therefore the doubling of the beats.

$$\text{Phase I, } y = \sin 2x + \sin 3x$$

$$\text{Phase II, } y = \sin(2x + 30) + \sin 3x$$

Figs. 3 and 4 show, from above downwards, with approximation and great simplification the successive positions of the phragma piece involved,

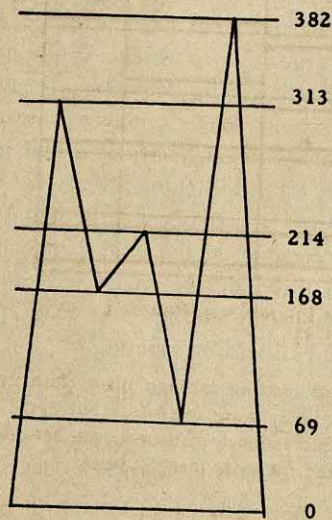


FIG. 1. DATA FOR PHASES I AND III
(Ratio 2:3)

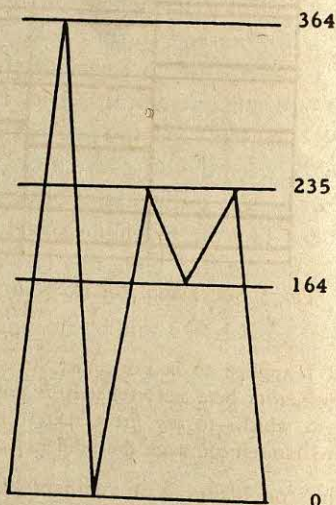


FIG. 2. DATA FOR PHASES II AND IV
(Ratio 2:3)

if the phragma is regarded as having the same width throughout and the hyperthesis is given no attention. In Fig. 4 we notice 71 units serving to throw 3 charges per time unit into the bundle of auditory nerve fibers located in the same neighborhood.

What about the (in the direction of the apex) following 164 units, from point 71 in Fig. 4 to point 235? We count there two charges, but in this Figure they do not look very evenly distributed in time. One time-interval looks smaller than the other; or in reverse, one time-interval looks larger than the other.

The last question raises a mathematic-mechanical problem which is beyond my mathematical ability. Surely the phragma units from point 71 to point 235 do not jump down all at one moment and up all at one moment. When I look at the action

of my hydraulic model of the cochlea I see it clearly that every point *has its own* jumping moment not shared by any other point, but the analytical mathematics, which should take the place of the graphics of the figures, is beyond me. It would be within my capabilities if it were a problem of wave mechanics (vibratory me-

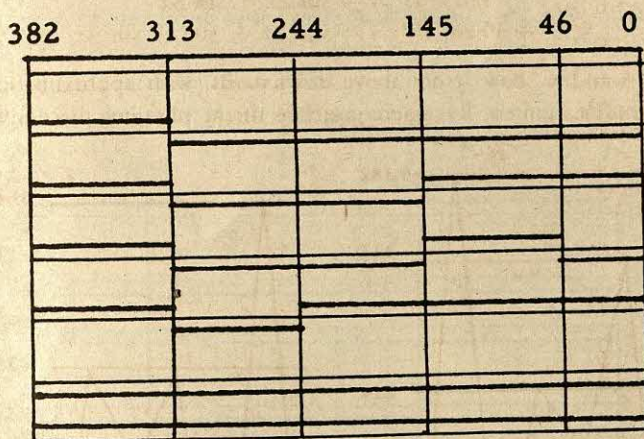


FIG. 3. SUCCESSIVE POSITIONS OF THE PHRAGMA IN PHASES I AND III
(Ratio 2:3)

chanics) applied to bodies under stress and reacting through plain elasticity, but the *hydraulics* here unfortunately is not ordinary wave mechanics but opens questions to which—to my great regret—no professional mathematician has hitherto applied himself although the need has existed for more than 50 years.

The conclusion is unavoidable that the regularity in time with which the two charges are sent brainwards from the points between 71 and 235

TABLE I
THEORETICAL LOUDNESS FOR THE PITCHES, RATIO 2:3

	Phase I	Phase II	Phase III	Phase IV
Pitch 3	46	71	46	71
Pitch 2	198	164	198	164
Pitch 1	138	129	138	129

is far greater than the poor approximation shown in Fig. 4 The mathematics determining how much regularity or irregularity there is in reality, remains to be solved.

Table I asserts that from the hydraulic theory the highest pitch must be expected noticeably to beat once in every half period. The beats must be expected to be doubled. The experimental observations mentioned above are in agreement with this prediction. The vibratory theories, to which

the mathematicians have given enormous time and effort, do not even try to relate their statements to the reports of experimental listeners. Why not?

(2) *Ratio 3:4*. A sinusoidal air wave was produced of a frequency of 480~. While this pitch is coming from the first generator, the second generator is tuned exactly giving, in musical terminology, the just interval of the fourth. This is left sounding, but the lower pitch is temporarily

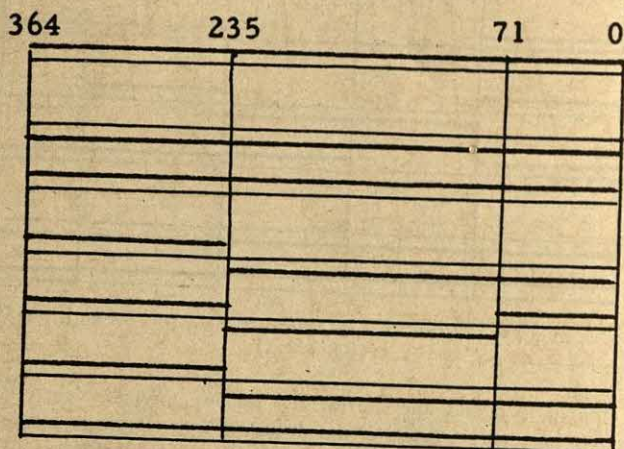


FIG. 4. SUCCESSIVE POSITIONS OF THE PHRAGMA IN PHASES II AND IV
(Ratio 2:3)

eliminated by using the volume control. The third generator is now brought in and tuned in unison to give about 3 beats in every 2 sec.

All three generators are now sounded together at a loudness a little below mezzoforte. Beats of the higher pitch are strongly heard. We turn the first generator off by use of the volume control, and the same beats are heard at the same rate. We bring the first generator back, and we hear the same beats.

Now we turn the second and justly tuned generator off through its volume control. We expect to hear beats again. Surprisingly, *no beats are heard*. It makes no difference whether we combine the Tone 3 with the justly tuned Tone 4 or with the mistuned Tone 4. In neither case are any beats of the Pitch 4 audible. Thus the experience is remarkably different from that we had with the ratio 2:3.

We have to ask: What does the hydraulic theory of cochlear mechanics make us expect?

Figs. 5 and 6 give us the computed ordinate differences of the maxima and minima of the following two equations. Phase III, gives the same results as Phase I, and Phase IV the same as Phase II. That is the reason why we expected to hear not single beats per period but double beats.

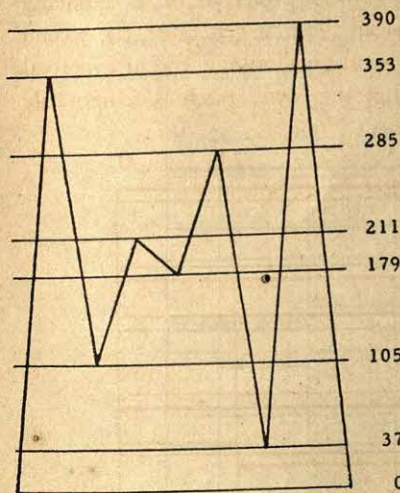


FIG. 5. DATA FOR PHASES I AND III
(Ratio 3:4).

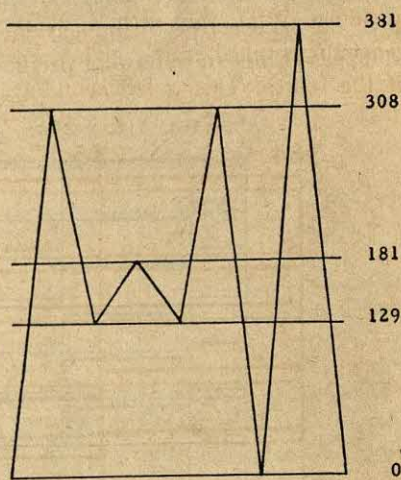


FIG. 6. DATA FOR PHASES II AND IV
(Ratio 3:4)

We hear neither. Let us proceed with applying the hydraulic theory, within the limits of our mathematical abilities.

$$\text{Phase I, } y = \sin 3x + \sin 4x$$

$$\text{Phase II, } y = \sin 3x + \sin(4x - 30)$$

Figs. 7 and 8 show, from above downwards, with approximation and great simplification the successive positions of the phragma points (not implying that they reach to the apex). In Fig. 8 we notice 52 units clearly serving to throw four charges per time unit into the neighboring *bundle* of auditory nerve fibers. Not so simple is the question concerning the result coming from the phragma region after point 52 and extending to point 179. Although it is very approximate, the description furnished by Fig. 8 seems to tell us that *three* charges are delivered. Are they, however, delivered with *sufficient or insufficient regularity* in temporal dispersal that through mediation of the nervous system the listener may report to hear the Pitch 3? This brings us back to the fact already discussed that no mathematician has applied himself to the *hydraulic* problem, hence we cannot answer the last question.

the time unit. From the experimental evidence (although the mathematical evidence is lacking) we should conclude that the latter is the truth. Unfortunately, however, I am *unable to separate* in Table II the loudness of the musical note b, Pitch 3, from the loudness of the next higher note e, Pitch 4, as inspection of the table will show.

Now from point 180 to point 316 in Fig. 7, and from point 179 to point 308 in Fig. 8, *two* charges are indicated with a long interval free from charges. This is as if we stopped all but two successive ones of the 48 holes of the siren disk men-

TABLE II
THEORETICAL LOUDNESSES FOR THE PITCHES, RATIO 3:4

	Phase I	Phase II	Phase III	Phase IV
Pitch 4 and Pitch 3 together	180	179	180	179
Pitch 1	210	202	210	202

tioned above. There is no certain evidence from siren-experiments that these two puffs alone are capable of giving the listener a definite perception of pitch. The two holes are like a single siren hole.

All that is logically possible to expect from the two charges together is to convey to the listener the Pitch 1, to be added to the Pitch 1 from the most extreme phragma units. Table II gives the summary of this the-

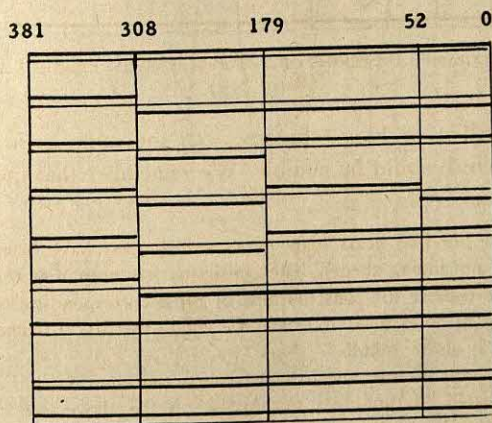


FIG. 8. SUCCESSIVE POSITIONS OF THE PHRAGMA IN PHASES II AND IV
(Ratio 3:4)

oretical discussion. The (musical) interval of the Fourth is the most confusing one both in experimenting and in theorizing, has always been that in pure audition and also in the esthetic theory of music.

There are no beats to be expected from the table;⁶ and none were heard. Not the faintest trace of a *summation* tone, Pitch 7, was ever audible.

(3) *Ratio 4:5*. A sinusoidal air wave was produced of a frequency around 300~. While this pitch is heard from the first generator, the second generator is tuned exactly, giving, in musical terminology, the just interval of the Major Third. This is left sounding, but the lower pitch is temporarily eliminated by using the volume control. The third generator is then brought in and tuned in unison to give about 3 beats in every 2 sec.

All three generators are now sounded at a loudness a little below mezzoforte. Beats of the higher pitch are strongly heard. We turn the first

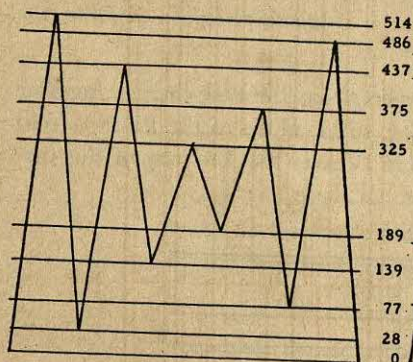


FIG. 9. DATA FOR PHASES I AND III
(Ratio 4:5)

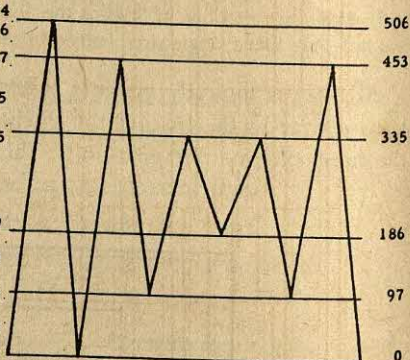


FIG. 10. DATA FOR PHASES II AND IV
(Ratio 4:5)

generator off by use of the volume control, and the beats are heard at the same rate. We bring the first generator back.

Now we turn the second and justly tuned generator off through the use of its volume control. We expect to hear beats again but surprisingly no beats are heard. It makes no difference easily describable whether we combine Tone 4 with the justly tuned Tone 5 or with the mistuned Tone 5. In neither case are beats of Pitch 5 clearly perceptible. What does the hydraulic theory of cochlear mechanics say to that?

Figs. 9 and 10 give us the computed ordinate differences of the maxima and minima of the following two equations. Phase III gives the same results as Phase I, and Phase IV the same as Phase II. Consequently we

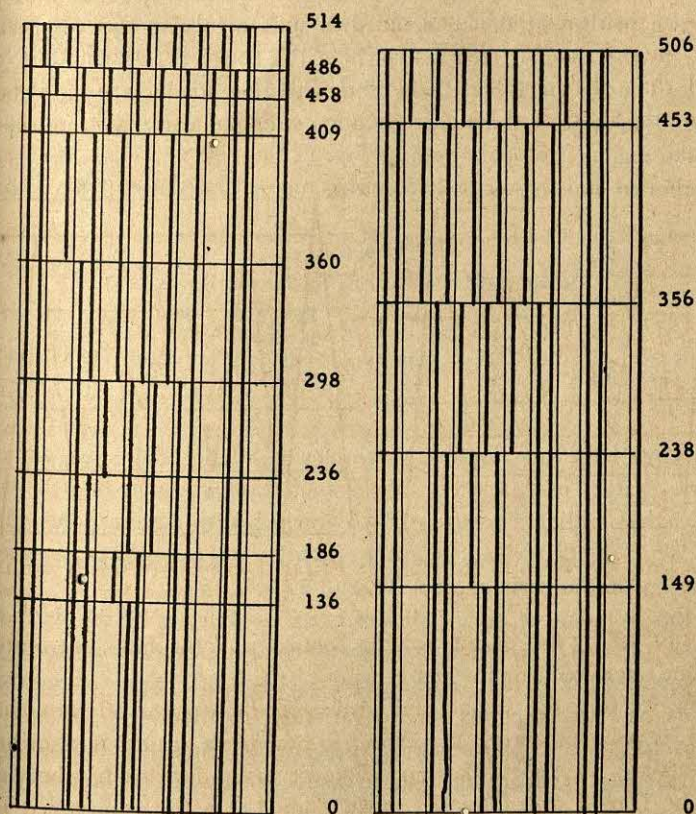
⁶ See footnote 7.

expected to hear even *doubled* beats, but we heard no beats at all. Let us proceed with applying the hydraulic theory (within our mathematical abilities).

$$\text{Phase I, } y = \sin 4x + 1.6 \sin 5x$$

$$\text{Phase II, } y = \sin(4x + 18) + 1.6 \sin 5x$$

Figs. 11 and 12 show, from above downwards, with approximation and great simplification the successive positions of the phragma points (not



implying that they reach to the apex). In Fig. 12 the first 149 units clearly serve to throw five charges per time unit into the neighboring *bundle* of nerve fibers. Not so simple is the question concerning the result coming

from the region after point 149 and extending to point 238. This brings us back to the already discussed lack of help from the mathematicians. If the *regularity of dispersal* suffices for the demands of the auditory nerve, then the *four* charges indicated by the figure must make the listener hear Pitch 4. A further question is presented by the region from point 238 to point 356 of Fig. 12. *Three* charges are indicated. If we regard a siren-experiment as applicable in which *every fourth hole* is stopped, then many of these 118 units would give an addition in loudness to the Pitch 4 just mentioned above; but *some of the more distant* of these 118 might be shown mathematically to occur with such regularity that they would convey to the listener Pitch 3, which in the above experiment is *weakly* audible. Thus in the Table III the Pitch 4 and the Pitch 3 are not assigned separate loudnesses because the separation would be guesswork.

In the region from point 356 of Fig. 12 to point 453 *only two* charges are indicated and these closely together and separated from the next two

TABLE III
THEORETICAL LOUDNESSES FOR THE PITCHES, RATIO 4:5

	Phase I	Phase II	Phase III	Phase IV
Pitch 5	149	136	149	136
Pitch 4 and Pitch 3 together	207	224	207	224
Pitch 1	150	154	150	154

charges by a large time interval. Judging by the results of siren-experiments, the two charges in such a case act as a single charge sent into the nerve bundle, thus adding Pitch 1 to Pitch 1 which comes from point 453 to point 506. Table III gives a summary of what seems to be the most *probable* conclusion on the basis of Figs. 11 and 12.

In the abstract, the table indicates beats for Pitch 5 by the fluctuation between 149 and 136, but in the experimental observations no such beats were reported. Why not? *There exists a threshold* for the consciousness of beats, as Fletcher points out.⁷ A variation from loudness 136 to 149 is only as 100 to 109. This seems to be below the threshold for consciously perceiving beats. (Warning: The hydraulic theory of cochlear *mechanics*

⁷ Harvey Fletcher (*Speech and Hearing*, 1929, 148), reporting the results of R. R. Riesz, writes: "The observer signaled that he heard beats. . . . After about twenty such judgments the operator was able to locate with considerable certainty the setting for which the observer was just able to detect a fluctuation in intensity. Below it the fluctuation was imperceptible. . . . For sensation levels above 50 db. the fractional change in the intensity which is just perceptible is between 5 and 10%. At a sensation level of 10 db. the fractional increase must be 73% to be perceptible."

is totally unrelated to the fact of a *threshold* of beat perception. The threshold for consciousness is the result of conditions existing in the *nervous system*.⁸ The hydraulic theory is concerned only with the mechanical frequency-analysis.)

The following important fact must again be emphasized. We never could hear in the experiments the faintest trace of a summation tone, *i.e.* Pitch 9. The common assertion by the textbooks that the ear produces *summation tones* is clearly fiction.

Exactly as with the ratio 4:5 I have performed experiments with the ratios 5:6 and 6:7. No audible beats resulted from slight mistuning. Obviously the beat threshold was not reached. And the only 'secondary' pitch which was definitely discernible was the Pitch 1. Especially this fact must again be emphasized in defiance of the common doctrine of the textbooks. Not the slightest trace of a *summation tone*, Pitch 11, was audible for the ratio 5:6; not the slightest trace of a *summation tone*, Pitch 13, was audible for the ratio 6:7.

A reader may be curious enough to ask how I succeeded in tuning the just interval 4:5 in spite of the fact that there are *no direct beats* audible when 4:5 is slightly mistuned. The answer is that loudspeakers produce some very high harmonics which are so faint that they escape ordinary detection, but become noticeable, though barely, when they are slowly beating. Now a somewhat musically gifted person can tune a Major Third with fair (though not absolute) accuracy merely by its melodiousness. Then one finds two positions of the vernier on both sides of a middle point which give those high pitches beating at the same slow rate. Setting the vernier then at that middle point, one eliminates those high beats and has secured the perfectly tuned Major Third.

(4) *Ratio 1:4*. Because of the immense possibilities of illusions in properly distinguishing and locating triple pitches in single and double Octave intervals *simultaneously* perceived, I devised the following experiment.

The first generator is tuned at 100~ and the second generator at 200~, all beats being eliminated. The first generator is then silenced and the third generator is tuned at 400~, all beats with the second generator's 200~ being eliminated. Now the second generator is tuned up an Octave,

⁸ In my book, *How We Hear* (Chapter XIII, B), when discussing multiple frequency beats, I failed to mention that all kinds of beats (and various kinds variously) depend for their audibility on more than mechanical production in the cochlea. They must also be above the neural threshold for perceptibility of the various kinds of beats.

to be in exact unison with the third generator, no beats being allowed. Then two positions of the *vernier* of the third generator ($400\sim$) are marked *on both sides* of its last position in order to give with the second generator ($400\sim$) *the same* rate of beats, U ($=$ unmeasured as absolute beat rate but obviously equal to the *physical* amount of mistuning, $400 + U$ or $400 - U$).

Now the first generator's $100\sim$ is sounded with the third generator's

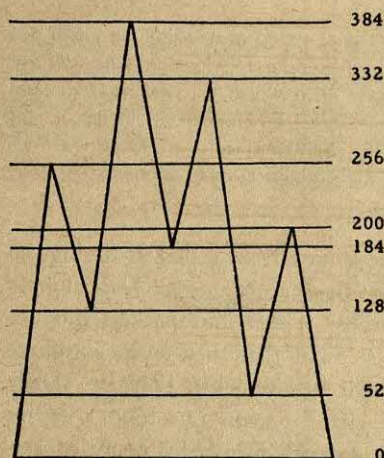


FIG. 13. DATA FOR PHASES I AND III
(Ratio 1:4)

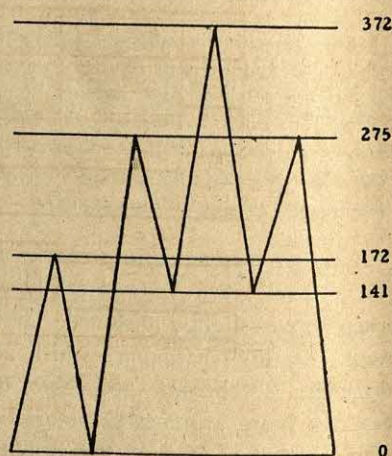


FIG. 14. DATA FOR PHASES II AND IV
(Ratio 1:4)

$400 + U$. There are no clear beats of either pitch 100 or pitch 400, but a 'secondary' pitch 200 beats at the rate of $2U$ per sec. Also the first generator's $100\sim$ is sounded with the third generator's $400 - U$. There are no clear beats of either Pitch 100 or Pitch 400, but the 'secondary' Pitch 200 beats at the rate of $2U$ per sec.

Figs. 13 and 14 give the computed ordinate differences of the maxima and minima of the following two equations. Phase III, gives the same result as Phase I, and Phase IV the same as Phase II.

$$\text{Phase I, } y = \sin x + \sin 4x$$

$$\text{Phase II, } y = \sin x + \sin (4x + 90)$$

Figs. 15 and 16 show, from above downwards, the approximate positions of the successive phragma points. Fig. 16 indicates that the loudness of the Pitch 4 should be relatively measured by the value 134.

In Fig. 15 the first 128 phragma units give us logically the Pitch 4. The

following 20 units present us with the case above discussed where every fourth hole of a siren-disk is stopped. So we must add to the 128 these 20 as also giving us the Pitch 4.

In Fig. 16 the 38 units from point 134 to point 172 indicate two charges sent to the neighboring bundle of nerve fibers. These charges look

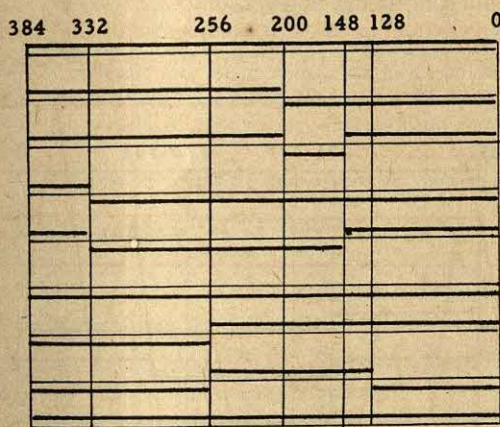


FIG. 15. SUCCESSIVE POSITIONS OF THE PHRAGMA IN PHASES I AND III
(Ratio 1:4)

in this Figure as if they occurred irregularly, but in the absence of an adequate mathematical analysis of the hydraulic functioning the Figure may deceive us. The two charges may occur with sufficient regularity. If they do, we should expect Pitch 2, hence I show it in Table IV. The next 200 units of Fig. 16 indicate Pitch 1.

If Table IV is the correct interpretation of our theoretical assumptions

TABLE IV

THEORETICAL LOUDNESSES FOR THE PITCHES, RATIO 1:4

	Phase I	Phase II	Phase III	Phase IV
Pitch 4	148	134	148	134
Pitch 2 (?)	zero	38	zero	38
Pitch 1	236	200	236	200

(I present it undogmatically,) the beats of Pitch 4 are not likely to surpass the beat threshold of perception: 134 to 148 is only as 100 to 110. Beats of Pitch 1 are also near the threshold, since 200 to 236 is as 100 to 118, and the sensation level in this Bass region is very low.⁹ Beats of Pitch 2 might

⁹ See footnote 7.

be theoretically allowed to be perceptible, for 0 to 38 (while being a small loudness) indicates a great loudness *fluctuation*. The experimentally observed doubling of the beats of Pitch 2 also appears in the table. Nobody can regret more than I that the acoustical literature contains so few experimental observations using methods similar to those described here and with which I could compare my observations.

Let us return to the question whether it is logical to give 'secondary' pitches the names of 'difference and summation' tones and 'aural harmonics' all of which are allegedly produced by the cochlea. With the ratio 1:4 not a trace of *harmonic*

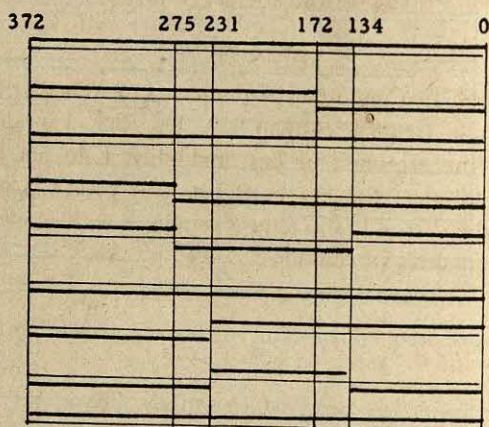


FIG. 16. SUCCESSIVE POSITIONS OF THE PHRAGMA IN PHASES II AND IV (Ratio 1:4)

Pitch 8, 12, 16, etc. was audible; not a trace of a *difference tone*, Pitch 3, was audible; not a trace of a *summation tone*, Pitch 5, was audible—and 2 is certainly not the difference between 1 and 4. The illogical nomenclature for secondary pitches ought to be removed promptly from the textbooks.

TABLE V
THEORETICAL LOUDNESSES FOR THE PITCHES, RATIO 2:5

	Phase I	Phase II	Phase III	Phase IV
Pitch 5	114	138	114	138
Pitch 2	214	184	214	184
Pitch 1	64	62	64	62

(5) *Ratio 2:5* In the same manner as in the cases described above, I experimented with the ratio 2:5. Pitch 2 was represented by the pure tone of 120~ and Pitch 5 by the pure tone of 300~. When 120~ and 300~ and $300 \pm U$ were all three sounded together, Pitch 300 was

heard beating at the rate of U per sec. When $300\sim$ was eliminated, and $120\sim$ with $300 \pm U\sim$ were sounded alone, the beats of the higher pitch were heard at the rate of $2U$ per sec. No *summation tone*, Pitch 7, was ever audible.

I made the theoretical calculation for the four characteristic phases of $y = \sin 2x + \sin 5x$. To save space, I give at once in tabular form the resulting loudness for the primary pitches and the one of Tartini.

Since 114:138 is equal to 100:121, the loudness fluctuations of Pitch 5 are obviously above the threshold for beats. The fluctuations occur twice during the period, in accord with the experimental observation.

CONCLUSIONS

I conclude with two important reminders. (1) Whether beats of other mistuned multiple frequencies than 1:2, 1:3, 2:3, 1:4 and 2:5 can be brought above the *threshold* of beat audibility, I do not know. (2) A *doubling* of multiple frequency beats for each period is mathematically possible only when an odd ratio term is combined with an even term, never when two odd numbers are combined as, *e.g.* in 1:3.

PROBABILITY-PREFERENCES AMONG BETS WITH DIFFERING EXPECTED VALUES

By WARD EDWARDS, The Johns Hopkins University

The decisions people make in gambling situations are psychologically interesting because they are motivated human responses and because they are primarily controlled by simple numerical properties of the stimulus-situation. In any situation in which a person must decide whether or not to bet on the occurrence of a future event, two sets of variables will probably be crucial in controlling his behavior: the probability of winning or losing, and the amounts which can be won or lost. If the amount of money (or any other measurable commodity) which can change hands for each possible outcome of a bet is multiplied by the probability of that outcome, the sum of these products over all possible outcomes is called the expected value (*EV*) of the bet. The *EV* is the amount the gambler should expect to win (or lose) on each play of the bet. If a bet has a positive *EV*, the gambler will in the long run profit by taking it. If it has a negative *EV*, he will in the long run lose by taking it. If it has an *EV* of zero, he is just as likely to win as to lose in the long run. This experiment is designed to identify variables which determine choices among bets which differ from one another in *EV*.

In a previous paper I reported an experiment in which bets of the same *EV* were compared with each other by the method of paired comparisons.¹ That experiment will hereafter be called the constant-*EV* experiment. Some bets were markedly preferred to others, even though, since all choices were

* Accepted for early publication July 16, 1953. This research was initiated and the experimental work was carried out at the Psychological Laboratories of Harvard University. It was supported in part by funds made available from the Laboratory of Social Relations, Harvard University, through its Project on the Nature of Risk and the Theory of Games. Further analysis of the data and preparation of this report were partially supported and done in cooperation with the Systems Coordination Division, Naval Research Laboratory under Contract N5-ori-166, Task Order I, between the Office of Naval Research and The Johns Hopkins University. This is Report No. 166-I-166, Project Designation No. NR-507-470, under that contract. This report is adapted from part of a doctoral dissertation submitted to Harvard University. The author is indebted to Professors J. G. Beebe-Center, G. A. Miller, and F. Mosteller, who contributed to the progress of this research.

¹ Ward Edwards, Probability-preferences in gambling, this JOURNAL, 66, 1953, 349-364. A fuller account of the methods, apparatus, and bets used in the experiment here being reported can be found in this paper.

between bets which had the same EV , there was no objective reason for preferring any bet to any other. For example, among positive EV and zero EV bets, those with a $4/8$ probability of winning were especially often chosen, and those with a $6/8$ probability of winning were especially seldom chosen. The results were attributed to the existence of consistent preferences for some probabilities over others. Another factor which might influence choices in gambling situations is the difference between subjective and objective values of money. Von Neumann and Morgenstern² have proposed and Mosteller and Nogee³ have employed a method for measuring the subjective value (which economists call the *utility*) of money by means of decisions in gambling situations. The results of the previous experiment raise doubts about this method. If choices among bets reflect preferences among probabilities as well as utilities of money, they cannot be used to measure either variable separately unless the other is controlled.

In the previous experiment, every choice was between two bets with the same EV . In this experiment, every choice is between two bets with different EV s. What should happen in this experiment? If such phenomena as preferences among probabilities and differences between objective and subjective values of money are ignored, we expect S to choose the bet with the larger EV (or the smaller negative EV , if he is required to accept a negative- EV bet. The failure of S to follow this simple EV -maximization strategy would suggest that something other than EV s was influencing his behavior, but would not show what. In order to show that probability-preferences were influencing his behavior, it would be desirable to be able to use previous knowledge of probability-preferences to predict deviations from simple EV -maximization. We will examine such deviations, and see whether they can be predicted from probability-preferences.

METHOD

Apparatus. The gambling in this experiment was done on the same pinball machine used in the constant- EV experiment. There are eight cells at the bottom of this machine, into any one of which the ball may roll, and its operation is surreptitiously controlled by 18 electromagnets, connected to an automatic programmer, which completely determine where the ball rolls. The S s are told that the ball is just as likely to roll into any one of the eight cells as into any other, and a randomly prepared tape is used in the programmer to make this statement true.

² J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, 1944. For a more complete discussion of this and other theories of utility measurement than this paper gives, see Edwards, *The theory of decision-making*, *Psychol. Bull.* (in press).

³ F. Mosteller and P. Nogee, An experimental measurement of utility, *J. Polit. Econ.*, 59, 1951, 371-404.

Bets. There were two parts to this experiment. On the first two days, bets were the same as those used in the constant-*EV* experiment. These bets are of three kinds. Eight of them have an *EV* of +\$0.525, and so will be called positive expected value, or *PEV*, bets. Eight of them have an *EV* of -\$0.525, and so will be called negative expected value, or *NEV*, bets. Eight of them have an *EV* of \$0.00, and so will be called zero expected value, or *ZEV*, bets. Each bet was paired with all others of its *EV*-level according to Ross's optimal order for pairs in the method of paired comparisons,⁴ and each pair was typed on a 3 x 5 card. There were thus 84 constant-*EV* cards.

In the last four days of this experiment, bets which differ from one another in *EV* were used. Table I reproduces these bets. Three criteria were followed in making up the differing positive *EV* (*DPEV*) bets in Table I, on which the *DNEV* and *DZEV* bets are based. First, a bet with a higher probability of winning was required to pay off less than a bet with a lower probability of winning, and no bet could involve administratively unreasonable amounts of money. Secondly, no two bets were permitted to differ by less than \$0.05 in actual amount of payoff, or by less than \$0.02 in *EV*. Thirdly, an attempt was made to choose bets with *EV*s such that consistent preference of the bet with the higher *EV* would result in a rank-ordering of the bets inverse to that found in the first two days with the constant-*EV* bets. These requirements turned out to be inconsistent with one another. In order to meet the first two, it was necessary to compromise the third. Consequently, the rank-order correlation (*rbo*) between the order of preferences (summed over *Ss*) found on the *PEV* cards in the first two days and the *EV* of the *DPEV* bets is only -0.76.

The *DNEV* bets were prepared by changing the verb in each *DPEV* bet from 'win' to 'lose.' Since the *DNEV* bets were not based on performance on *NEV* bets in the first two days, it is not surprising that the correlation between order of preferences on the *NEV* bets during the first two days and the *EV* of the *DNEV* bets is only -0.10.

The *DZEV* bets were prepared by combining each *DPEV* bet with a *DNEV* bet so that the probabilities involved summed to one. The correlation between *EV* of the *DZEV* bets and choice on the *ZEV* bets in the first two days is -0.89.

The mean difference in *EV* between all pairs of *DPEV* bets is slightly less than \$0.12. The same is of course true of *DNEV* bets. The corresponding figure for *DZEV* bets is \$0.09. This means that consistent choice of the alternative with the higher *EV* or lower negative *EV* should over the long run result in each *S* winning about \$9.24 per hr. No one did so.

The differing-*EV* bets also were typed on cards in pairs, so there were 84 differing-*EV* cards. The apparatus thus consisted of pinball machine, programmer and tape, a deck of constant-*EV* cards, a deck of differing-*EV* cards, and plenty of poker chips.

Subjects. The *Ss* were 8 men undergraduates from the Harvard Summer School. The names of those to be invited were randomly selected from the complete list of summer school students, but it was necessary to issue 16 invitations to get 8

⁴ R. T. Ross, Optimum orders for the presentation of pairs in the method of paired comparisons, *J. Educ. Psychol.*, 25, 1934, 375-382.

acceptances. To the extent to which refusal to participate in the experiment biases the sample, therefore, the Ss do not constitute a random sample. Even if the sample were representative of Harvard Summer School it would not be representative of Harvard College much less of undergraduates in general or of any larger population. The results will show that this state of affairs is not as unfortunate as it could have been.

TABLE I

BETS OF DIFFERING EXPECTED VALUE USED IN THIS EXPERIMENT

Bets with Positive Expected Value

- (1) If you roll a 4, you win \$6.00. If you roll anything else, you win nothing. ($EV = \$0.75$ per roll)
- (2) If you roll a 1 or a 7, you win \$2.90. If you roll anything else, you win nothing. ($EV = \$0.73$ per roll)
- (3) If you roll a 2, a 4, or a 6, you win \$1.75. If you roll anything else, you win nothing. ($EV = \$0.66$ per roll)
- (4) If you roll a 2, a 4, a 7, or an 8, you win \$1.20. If you roll anything else, you win nothing. ($EV = \$0.60$ per roll)
- (5) If you roll a 2, a 3, a 5, a 7, or an 8, you win \$1.10. If you roll anything else, you win nothing. ($EV = \$0.69$ per roll)
- (6) If you roll anything but a 3 or a 6, you win \$1.05. If you roll a 3 or a 6, you win nothing. ($EV = \$0.79$ per roll)
- (7) If you roll anything but a 5, you win \$1.00. If you roll a 5, you win nothing. ($EV = \$0.88$ per roll)
- (8) Regardless of what you roll, you win \$0.85. ($EV = \$0.85$ per roll)

Bets with Negative Expected Value

These eight bets are identical with those listed above except that the verb in each bet is 'lose' instead of 'win.' Consequently, the expected values are also the same except that they are expected losses rather than winnings.

Bets with Zero Expected Value

- (1) If you roll a 4, you win \$6.00. If you roll anything else, you lose \$1.00. ($EV = -\$0.13$ per roll)
- (2) If you roll a 1 or a 7, you win \$2.90. If you roll anything else, you lose \$1.05. ($EV = -\$0.06$ per roll)
- (3) If you roll a 2, a 4, or a 6, you win \$1.75. If you roll anything else, you lose \$1.10. ($EV = -\$0.03$ per roll)
- (4) If you roll a 2, a 4, a 7, or an 8, you win \$1.20. If you roll anything else, you lose \$1.20. ($EV = \$0.00$ per roll)
- (5) If you roll a 2, a 3, a 5, a 7, or an 8, you win \$1.10. If you roll anything else, you lose \$1.75. ($EV = \$0.03$ per roll)
- (6) If you roll anything but a 3 or a 6, you win \$1.05. If you roll a 3 or a 6, you lose \$2.90. ($EV = \$0.06$ per roll)
- (7) If you roll anything but a 5, you win \$1.00. If you roll a 5, you lose \$6.00. ($EV = \$0.13$ per roll)
- (8) Regardless of what you roll, you neither win nor lose. ($EV = \$0.00$ per roll)

Procedure. The procedure was the same as in the real-gambling sessions of the previous experiment. The Ss were told that they could win or lose plenty of money, and that if they lost, they would have to pay out of their own money, but that the situation was biased in their favor to the extent of \$1.00 a session—each S was given \$1.00 more in chips at the beginning of each session than he was required to return at the end. The Ss were run individually and never met one another.

The first two sessions used the constant-*EV* cards. The next four sessions used the differing-*EV* cards. All data reported here are based on these six sessions. There were also one or two post-experimental sessions, whose sole purpose was to insure that no one lost money in the experiment as a whole. The net effect of these extra sessions was that each *S* recovered the amount of his losses, if any, plus enough more to bring his net gain from the experiment up to \$1.00 per hr. This devious method of paying *Ss* off prevented the word from getting around the Harvard campus to potential *Ss* for future experiments that no one could lose in the psychology gambling experiments.

Since the constant-*EV* bets in this experiment exactly duplicate those of the constant-*EV* experiment, and were offered to *Ss* under exactly the same conditions except that there had not been six preliminary non-gambling sessions, the results of the first two days of this experiment offer some small check on the reliability of the probability-preferences previously discovered. The *Ss* are different and the preceding experiences are different, but all other conditions are the same.

RESULTS

Constant-EV bets. Table II shows the results for the *PEV*, *NEV*, and *ZEV* bets of this experiment and the corresponding results from the con-

TABLE II
RESULTS FOR THE CONSTANT-*EV* CARDS IN THE PRESENT EXPERIMENT AND IN
THE PREVIOUS EXPERIMENT

Probability*	PEV		NEV		ZEV	
	present	previous	present	previous	present	previous
1/8	0.1674	0.1116	0.1585	0.1957	0.1897	0.1399
2/8	0.1339	0.1265	0.1629	0.1719	0.1696	0.1592
3/8	0.1607	0.1376	0.1317	0.1652	0.1473	0.1562
4/8	0.1942	0.1830	0.0982	0.1116	0.1763	0.2009
5/8	0.1027	0.1406	0.1161	0.1049	0.1161	0.1339
6/8	0.0603	0.0900	0.1719	0.1265	0.0592	0.0677
7/8	0.0893	0.0975	0.1183	0.0871	0.0670	0.0804
8/8	0.0915	0.1131	0.0424	0.0372	0.0647	0.0618
	$r=0.76$		$r=0.83$		$r=0.90$	

* For *PEV* and *ZEV* bets, the probability of winning; for *NEV* bets, the probability of losing.

stant-*EV* experiment. Each number appearing in the table is an index of relative preference for a bet over all others. It is computed by counting the total number of choices by all *Ss* of a bet over all others with which it was compared and dividing by the total number of choices made by all *Ss* on all bets at that *EV*-level. The maximum value this number can have for any bet is 0.25, which can only occur if everyone chooses the bet every time he has the chance. The minimum is 0.00, which can only occur if the bet is never chosen.

The correlation coefficients included in Table II are the product-moment

correlations between the results of the two experiments. It is readily apparent that the results from the two experiments are quite similar. (The similarity is even more apparent if the numbers are plotted on a graph.) The differences in experimental procedure, in population sampled, in size of sample, and in satisfactoriness of sampling procedures between this and the constant-*EV* experiment evidently did not seriously influence the pattern of probability preferences found in the constant-*EV* experiment.

Differing-EV bets. Fig. 1 shows the results for the *DPEV* bets, compared with the *PEV* bets in the first two days. The Y-axis is the index of preference explained above, and the X-axis represents the probability of winning. These probabilities are arranged in order of decreasing *EV* for the *DPEV* bets, so that consistent choice of the bet with the higher *EV* would result in the downward-sloping straight line shown. Clearly there is a general trend downward from left to right, as the strategy of maximizing *EV* would require, but there are also marked deviations from that strategy. There is, however, no evidence that the deviations from maximization of *EV* are related to the probability-preferences found in the first two days.

In Fig. 2, which presents the results of the *DNEV* bets, consistent choice of the bet with the smaller negative *EV* would result in the straight ascending line shown. It is apparent that the deviations from this pattern observed at $8/8$, $1/8$, and $5/8$ are all associated with corresponding decreases in the probability-preferences. The same can be said for the flattening at $4/8$. In short, the deviations from maximization of *EV* here can be predicted quite well from probability-preferences.

The fact that in the *DNEV* bets there is essentially no correlation between prediction based on probability-preference and prediction based on *EV*-maximization makes possible a very simple statistical test of whether or not probability-preferences predict better than chance in this situation. Since each choice is between two alternatives, and probability-preferences predict the selection of one of these alternatives, on the null hypothesis this prediction would be right 50% of the time. A simple chi-square test shows that this null hypothesis can be rejected beyond the 0.001-level of confidence when predictions are based on each individual's rank-ordering of the probabilities determined from his choices during the first two days. If the predictions are based on the communal probability preferences shown in Fig. 2, then the chi-square falls very slightly short of the 0.05-level.

Fig. 3 shows the *DZEV* results. Again *EV*-maximization would produce the descending line shown as with the *DPEV* graph, except that there is no

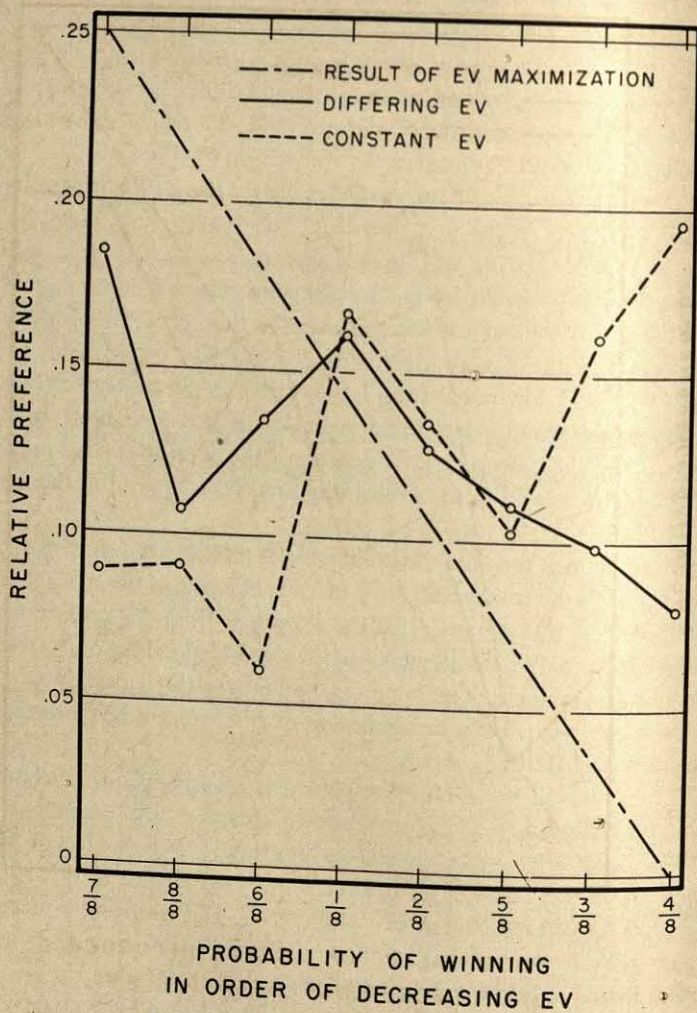


FIG. 1. RESULTS FOR THE DPEV BETS

The solid line shows the results for the DPEV bets; the dashed line shows the results from the PEV bets; and the dot-and-dashed line shows the results which would have been obtained if all Ss had followed a strategy of EV-maximization on the DPEV bets. The probabilities on the X-axis are arranged in order of decreasing EV.

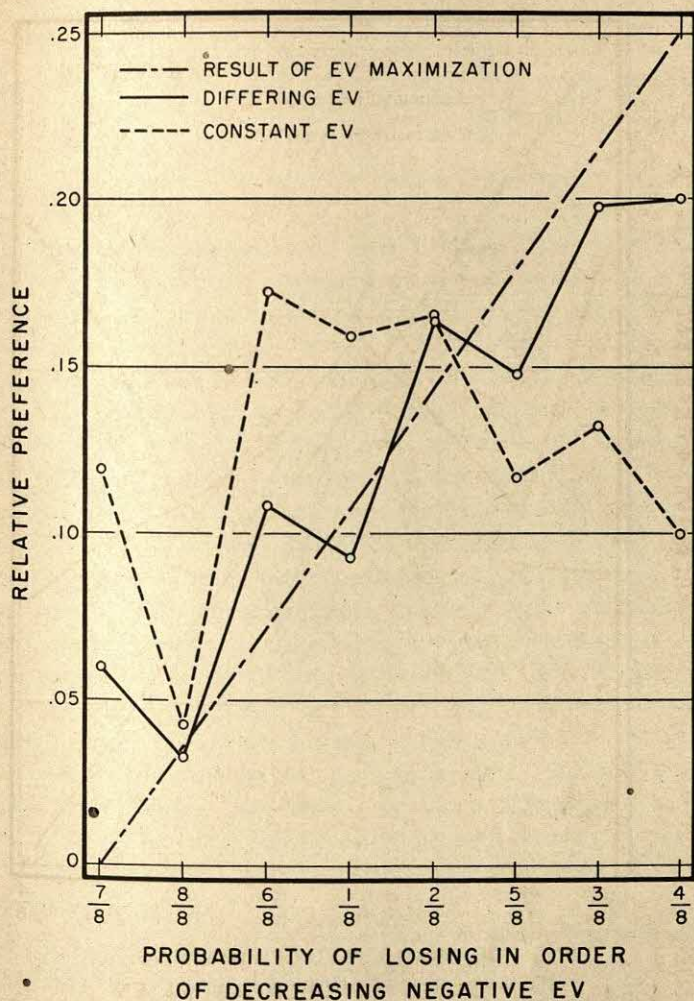


FIG. 2. RESULTS FOR THE DNEV BETS

The solid line shows the results from the DNEV bets; the dashed line shows the results from the NEV bets; and the dot-and-dashed line shows the results which would have been obtained if all Ss had followed a strategy of EV-maximization on the DNEV bets. The probabilities on the X-axis are arranged in order of decreasing negative EV.

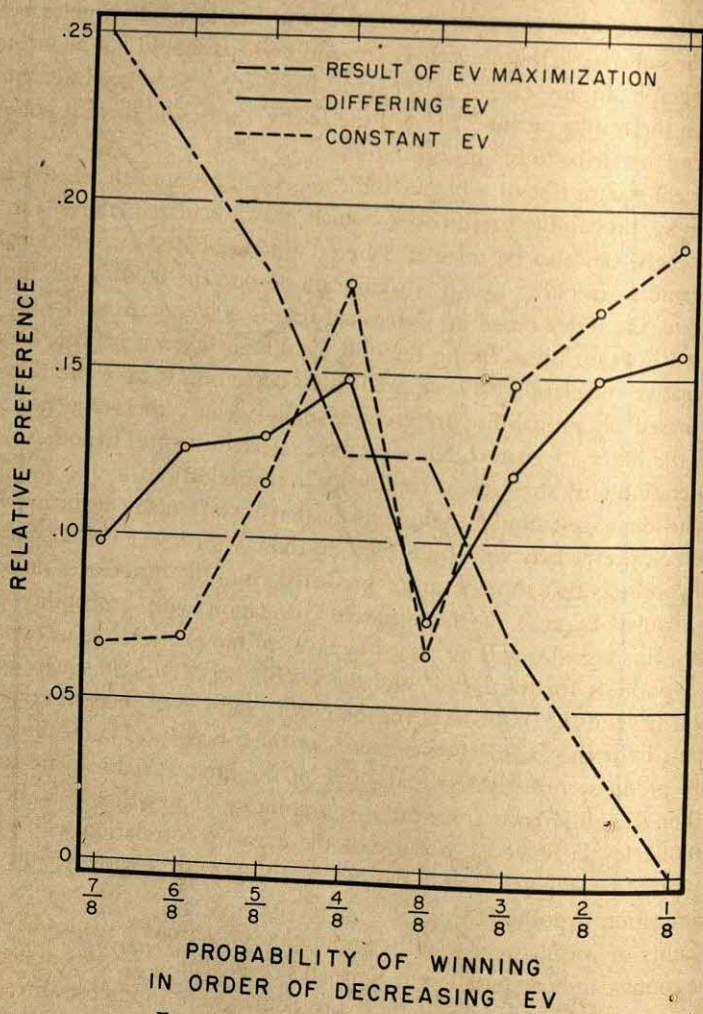


FIG. 3. RESULTS FOR THE DZEV BETS

The solid line shows the results from the DZEV bets; the dashed line shows the results from the ZEV bets; and the dot-and-dashed line shows the results which would have been obtained if all Ss had followed a strategy of EV-maximization on the DZEV bets. The probabilities on the X-axis are arranged in order of decreasing EV, except that the 4/8 and 8/8 probabilities of winning have the same EV.

difference in *EV* between the 4/8 bet and the 8/8 bet, so that *Ss* should have no basis for choice between these two. Actually, the 8/8 bet reads "regardless of what you roll, you neither win nor lose." It is not a bet in quite the same sense that the others are, and perhaps should not be included in the graph. In any case, it is apparent that there is very good agreement between the results on the *ZEV* and the *DZEV* cards, and there is not even any trend to ascribe to *EV*-maximization.

The chi-square test of whether the choices can be predicted better than chance by probability-preferences, which was calculated above for the *DNEV* bets, can also be calculated here. Whether based on individual or group results, in either case it is significant beyond the 0.001 level, though of course, the index based on individual results is larger. Actually, the test is biased in favor of accepting the null hypothesis when it is false, because the negative correlation between predictions based on *EV*-maximization and those based on probability-preferences should reduce the frequency with which the latter are correct. Since we have rejected the null hypothesis, this consideration only strengthens confidence in that rejection.

A chi-square test similar to those made above can be made of the hypothesis that *Ss* prefer bets with higher *EV* to those with lower *EV*, against the null hypothesis that choices can be predicted from *EV*-differences no better than chance. Because of the negative correlation with probability preferences, this test also will be biased in favor of too many acceptances of the null hypothesis for the *DPEV* and the *DZEV* bets. This chi-square is significant beyond the 0.001 level for the *DPEV* and *DNEV* bets. It cannot be computed for the *DZEV* bets because for these bets the *EV*-maximization theory predicts correctly less than 50% of the time. (A chi-square test of whether this difference from 50% is significant shows that it is, beyond the 0.001 level; however, because of the negative correlation with probability-preferences such a test has no meaning with respect to the *EV*-maximization hypothesis.)

Counts of inconsistent triads, computation of various statistics based on these counts, and computation of Kendall's coefficient of agreement between *Ss* were made.⁵ The results of these procedures were in almost all respects similar to those found in the constant-*EV* experiment, and so will not be reported here. The only discernible difference is that there seems to be a decrease in consistency as the experiment progresses from *PEV* to *DPEV* bets, rather than the usual increase; but this is probably a random variation.

⁵ M. G. Kendall, *Rank Correlation Methods*, 1948.

DISCUSSION

It certainly is clear from the results shown above that *Ss* do not consistently prefer bets with higher *EVs* to bets with lower *EVs*. Part of the variation from this simple strategy could have been predicted by probability-preferences and part could not. The *DPEV* bets, in which prediction from probability-preferences was least satisfactory, were the bets in which probability-preferences and *EV*-differences were strongly in opposition to each other; one or the other had to emerge as the major determinant of choices, and in this case the *EV*-differences turned out to be the more important. In the *DNEV* bets, in which the two determinants were essentially unrelated to one another, a compromise took place; both determinants were effective. In the *DZEV* case, once again the two determinants were strongly opposed to each other, and this time probability-preferences dominated—probably because the *DZEV* bets were more complicated, so that *EV*-differences were much more difficult to detect, and the *EV*-differences themselves were somewhat smaller.

It is reasonable to suppose that as the difference in *EV* between two bets, of which *S* must choose one, increases, it becomes more and more important in determining *S*'s choice, and probability-preferences therefore become less important. It is quite clear, however, at least within the range of *EV* differences studied in this experiment, that probability-preferences do partially determine choices among bets even though they differ in *EV*. This means that the difficulty for von Neumann and Morgenstern's utility-measurement proposal which was discussed in the introduction of this report (and more fully in the report of the constant-*EV* experiment) remains a difficulty. Any experiment designed to measure utility of money by means of gambling decisions must operate in the range of small *EV*-differences where this experiment shows probability-preferences are important.

It is unfortunate that probability-preferences and differences in expected value or in expected utility both determine decisions in gambling situations. If expected utility were the only important variable, the wonderful mathematical model provided by utility theory would predict such decisions in great detail. Even so, it might eventually be possible to develop separately subjective scales of probability (which would allow for probability-preferences) and of utility. Then the mathematical model could perhaps be used in conjunction with these scales to predict risky decisions. Here is a problem for much future research.

SUMMARY

Preferences were found for some probabilities over others in gambling situations in a previous experiment. In that experiment, all choices had to be made between bets of the same expected value. The present experiment was designed to find out what Ss would do when required to choose between bets of different expected values. After preliminary sessions in which the results of the previous experiment were confirmed, Ss were required to make choices between pairs of bets of different expected value according to the method of paired comparisons. The results showed that both preferences for bets with higher expected values, or lower negative expected values, and preferences among probabilities influenced S's choices. The former factor was dominant in those bets on which Ss could only win or break even; both were important in those bets on which Ss could only lose or break even; and the latter factor was dominant in those bets on which Ss could either win or lose. These variations in effectiveness of the two major determinants of choices were related to specific properties of the bets involved.

It is concluded that probability-preferences are important in determining decisions among bets even when there are objective reasons for preferring one bet to another, although as the difference in expected value increases it seems likely that the importance of probability-preferences will decrease.

THE RELIABILITY OF PROBABILITY-PREFERENCES

By WARD EDWARDS, The Johns Hopkins University

Two major properties of any risk-situation which may influence behavior are the possible outcomes of the risk and the probabilities of these outcomes. These two variables together determine the expected value (EV) of the risk. If the value of each possible outcome is measured in dollars or some similar unit, then the sum over all possible outcomes of the value of each outcome multiplied by its probability of occurrence is the EV . (For instance, I toss a coin. If it comes up heads, I pay you \$2.00. If it comes up tails, you pay me \$1.00. Then the $EV = 0.5 (-\$2.00) + 0.5 (+\$1.00) = -\$0.50$, which is the EV of the game from my point of view. From your point of view, it is of course $+\$0.50$.)

Two previous papers have reported that choices between bets of equal EV , presented in pairs to the Ss , are influenced by preferences among the probabilities involved,¹ and that this influence can be demonstrated to operate even when Ss are required to choose between bets with differing EVs .² These experiments raise a variety of methodological issues. In an area of human behavior so subject to individual differences as gambling, are these results reproducible? What are the effects of wording of the bets? Would similar results occur with different numbers of bets? What effect does the outcome of previous bets have on a decision? Would similar re-

* Accepted for early publication August 26, 1953. This research was partially supported and done in cooperation with the Systems Coordination Division, Naval Research Laboratory, under Contract N5-ori-166, Task Order I, between the Office of Naval Research and The Johns Hopkins University. This is Report No. 166-I-173, Project Designation No. NR-507-470, under that contract. Experiment II and part of Experiment I in this report were carried out in the Psychological Laboratories of Harvard University and were partially supported by funds made available from the Laboratory of Social Relations, Harvard University, through its Project on the Nature of Risk and the Theory of Games. This report of those experiments is adapted from part of a doctoral dissertation submitted to Harvard University. I am grateful to Mrs. Harriott Quin for collecting and processing data in those experiments which were performed at The Johns Hopkins University.

¹ Ward Edwards, Probability-preferences in gambling, this JOURNAL, 66, 1953, 349-364. In the rest of this paper, this first experiment in the series will be called the original experiment or the constant- EV experiment, and will not be separately cited.

² Edwards, Probability-preferences among bets with differing expected values, this JOURNAL, 67, 1954, 56-67. For a discussion of the significance of these experiments for the general theory of decision-making, see Edwards, The theory of decision-making, *Psychol. Bull.*, in press.

sults occur if the *EV*-levels chosen, and therefore the amounts of money involved in the bets, were different? Can probability-preferences be demonstrated in non-gambling situations? This paper reports a series of experiments which provide some answers (however incomplete) to these questions.

METHOD

Apparatus. The gambling in almost all of these experiments was done on a rigged pinball machine described in the first paper of the series. There are eight cells at the bottom of the machine into any one of which the ball may roll. The *Ss* are told that the ball is just as likely to roll into any one cell as into any other. For the previously reported experiments, that was true. The pinball machine is controlled by electromagnets and a tape-reader, and in previous experiments the tape was prepared from a table of random numbers. In some of the experiments reported here, the same procedure was used, but in others special rolls on particular bets were necessary, and the tape was prepared accordingly. Unless it is specified otherwise in the following experiments, the tape was random.

Bets and subjects. The bets and *Ss* will be described below separately for each experiment.

Procedure. In those experiments in which *Ss* were run individually, they were first shown the pinball machine, told about the bets, and told that they may either win or lose, but that the situation was biased in their favor by \$1 per session. This bias was achieved by giving each *S* \$21 in chips at the beginning of each session and requiring him to return only \$20 in chips or in cash at the end of the session. The *Ss* knew that if they had less than \$20 in chips, they must pay the difference in cash from their own funds. In those experiments in which money was not involved, *Ss* were paid for their time at the end of each session. As in the previously reported experiments, *Ss* were run in extra sessions (after the experimental ones) in which *E* made sure that they won back the amount of their losses, if any, plus enough more to pay for their time at the rate of \$1 per hour.

EXPERIMENT I: RELIABILITY

The choices of two different groups of *Ss* were studied. One group of *Ss* was put through a repetition of the first experiment in which bets of the same *EV* were compared with one another. This repetition of the constant-*EV* experiment was actually intended not only to test reliability, but also to establish baseline behavior for the experiment with the military game to be described later in this report (Experiment III). The second group of *Ss* was a classroom group. Slides of the same pairs of bets as were used in the constant-*EV* experiment were shown to them, and they were asked to imagine they were gambling and make choices. The choices of both these groups of *Ss* were compared with the corresponding choices of *Ss* in the original constant-*EV* experiment.

Bets. There were 24 bets. Eight of them, with probabilities of winning running

from $1/8$ through $8/8$, had an *EV* of $+\$0.525$ per roll, and so will be called the positive *EV* (*PEV*) bets. Eight of them had an *EV* of $-\$0.525$ per roll, and so will be called the negative *EV* (*NEV*) bets. Eight of them had an *EV* of $\$0.00$, and so will be called the zero *EV* (*ZEV*) bets. All bets of a given *EV* were paired with one another according to Ross's optimal order for pairs in the method of paired comparisons.³ Thus there were 84 pairs of bets in all. (For a fuller description of the bets, see the report of the constant-*EV* experiment, or see Table VIII.)

Subjects. The *Ss* for the full repetition of the constant-*EV* experiment (which will hereafter be called the replication) were 11 men, randomly chosen Johns Hopkins undergraduates. The *Ss* in the classroom-group were 47 members of a class in educational psychology in the Summer School of Education at Harvard. The members of this class were mostly high school teachers seeking advanced degrees, and so were very different both from the Hopkins undergraduates and from the Harvard undergraduates used in the original constant-*EV* experiment.

Procedure. The procedure used in the replication was the same as that used in the original experiment, except that only in the first two sessions (instead of the first four) were *Ss* run under the condition of just imagining they were gambling, and only the next two sessions (instead of the next six) involved gambling for chips which were not exchanged for money. The last four sessions in both experiments involved real gambling. (Another difference in procedure was that a research assistant collected the data in this experiment, while the author collected the data in the previous experiments.)

In the classroom-situation, a slide of the pinball machine was shown the students, and the general nature of the experiment was explained. Then the slides containing pairs of bets were presented, one by one. They were presented in order, first *PEV*, then *NEV*, then *ZEV*. This is, of course, a procedure different from that used in the constant-*EV* experiment and the present replication of it, in which the *PEV*, *NEV*, and *ZEV* cards were shuffled. Unfortunately, the end of the hour was approaching when the *ZEV* bets were reached, and it was necessary to hurry the students in making their choices. This fact is reflected in the *ZEV* results.

Results: (1) Imaginal sessions. Fig. 1 shows the results of the *PEV* bets from the classroom situation and this replication of the constant-*EV* experiment for the just-imagining sessions. The *PEV* results of the constant-*EV* experiment just imagining sessions are included for comparison. The measure of relative preference on the ordinate is the number of choices for each bet over all others with which it is compared, divided by the total number of judgments made at this *EV*. It is at once apparent that the agreement is excellent among the three curves. Fig. 2 shows the same results for the *NEV* bets, and Fig. 3 for the *ZEV* bets. Again agreement is excellent. The flattening of the *ZEV* classroom results is caused by the speeding of judgments made necessary by the approaching end of the hour. The assertion that these curves appear similar is supported by the prod-

³ R. T. Ross, Optimum orders for the presentation of pairs in the method of paired comparisons, *J. Educ. Psychol.* 25, 1934, 375-382.

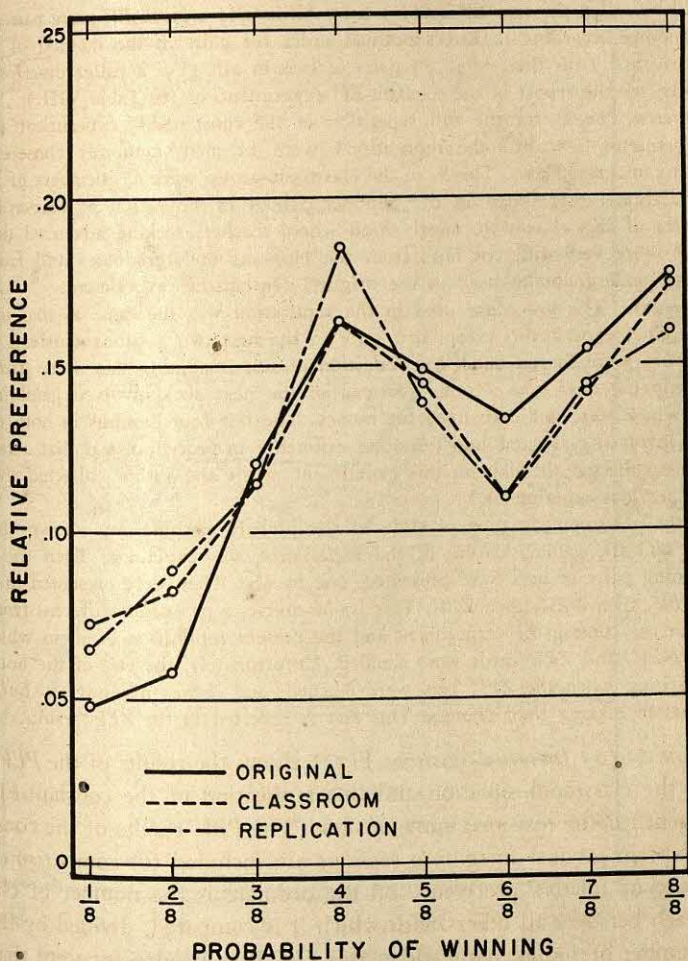


FIG. 1. PEV RESULTS FOR JUST-IMAGINING SESSIONS IN EXPERIMENT I
 The curve labelled "Replication" shows the results of this replication of the constant-EV experiment. The curve labelled "Classroom" shows the results of the classroom experiment. The curve labelled "Original" shows the PEV results for the constant-EV experiment just-imagining sessions, which are included for comparison.

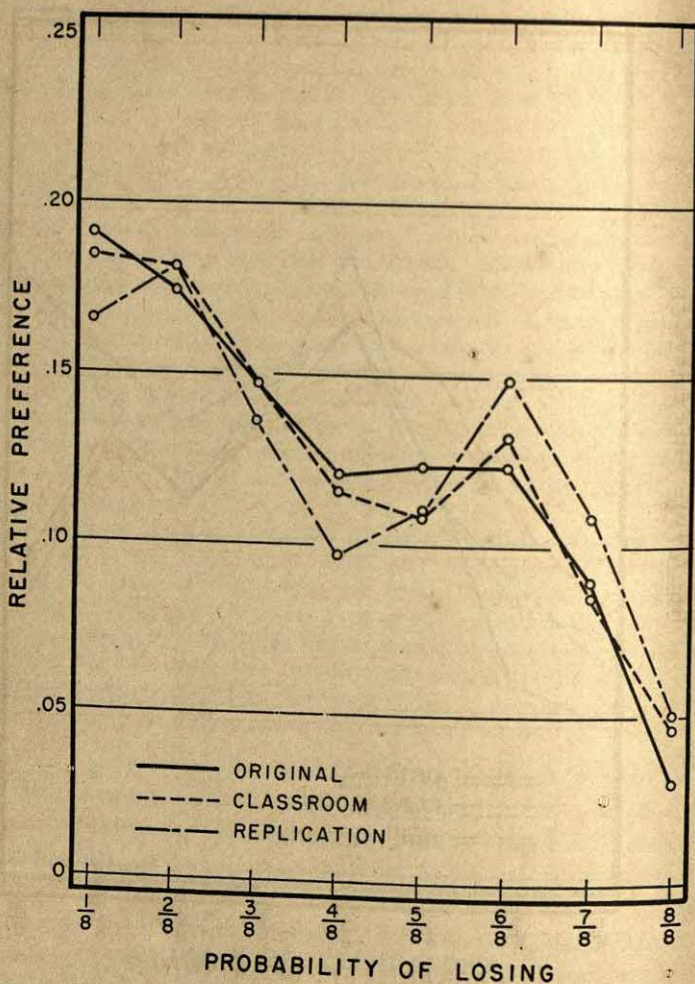


FIG. 2. NEV RESULTS FOR JUST-IMAGINING SESSIONS IN EXPERIMENT I
 The curve labelled "Replication" shows the results of this replication of the constant-EV experiment. The curve labelled "Classroom" shows the results of the constant-EV experiment just-imagining sessions, which are included for comparison. The curve labelled "Original" shows the PEV results for the

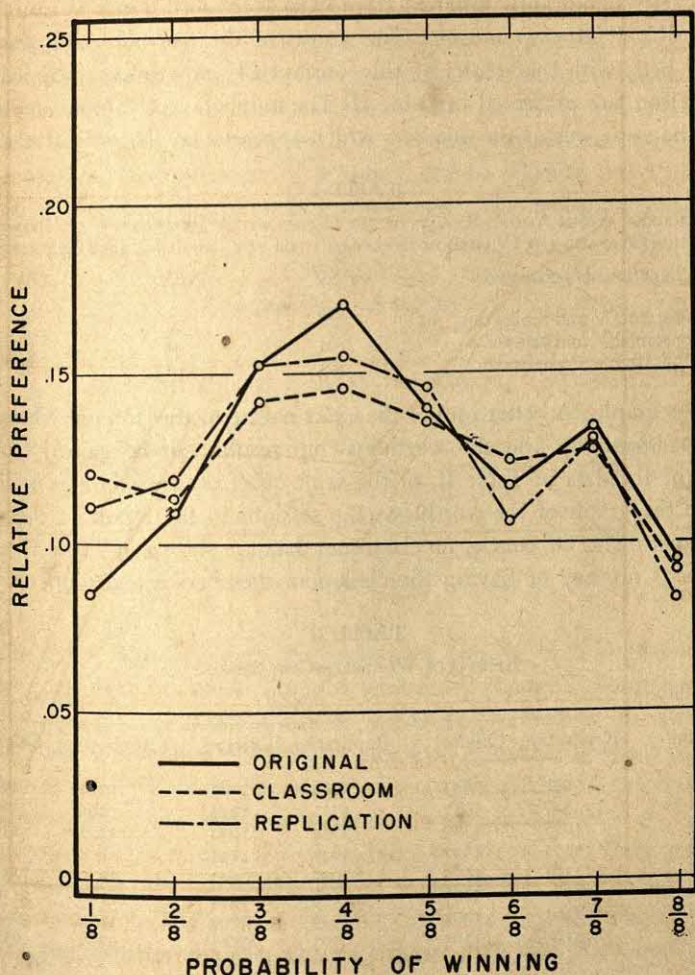


FIG. 3. ZEV RESULTS FOR JUST-IMAGINING SESSIONS IN EXPERIMENT I
 The curve labelled "Replication" shows the results of this replication of the constant-EV experiment. The curve labelled "Classroom" shows the results of the classroom experiment. The curve labelled "Original" shows the PEV results for the constant-EV experiment just-imagining sessions, which are included for comparison.

uct-moment correlation coefficients between them given in Table I. For these and for all other correlation coefficients to be reported in this paper, $r = 0.71$ is significantly different from Zero at the 0.05-level.

(2) *Worthless-chip sessions.* The results of the worthless-chip sessions for all bets, with the results of the constant-EV experiment included for comparison, are presented in Table II. The numbers in the table are calculated the same way as the numbers which appeared on the ordinates of the

TABLE I

INTERCORRELATIONS AMONG RESULTS OF THE CONSTANT-EV EXPERIMENT, THE PRESENT REPLICATION, AND THE CLASSROOM EXPERIMENT FOR THE JUST-IMAGINING CONDITION

Experiments correlated	PEV	NEV	ZEV
Constant-EV and replication	.88	.92	.78
Constant-EV and classroom	.95	.98	.85
Replication and classroom	.93	.95	.86

previous graphs. A better idea of the exact nature of the difference between the just-imagining and the worthless-chip results can be gained from a graph of the data in Table II, or the same effect can be achieved by looking up the graph of the worthless chip-sessions in the report of the original experiment. Of course, no classroom data are shown in Table II, since there was no way of having the classroom members actually operate the

TABLE II

RESULTS OF WORTHLESS-CHIP SESSIONS

Probability of winning or losing	PEV		NEV		ZEV	
	Replication	Original	Replication	Original	Replication	Original
1/8	.0958	.0833	.1964	.2183	.1201	.1091
2/8	.1347	.1037	.1753	.1944	.1299	.1265
3/8	.1055	.1329	.1737	.1632	.1201	.1384
4/8	.2175	.1895	.0844	.1195	.1801	.1984
5/8	.1185	.1305	.0990	.0942	.1429	.1354
6/8	.0909	.1096	.1299	.1091	.1136	.1106
7/8	.1201	.1200	.0909	.0699	.1169	.1240
8/8	.1169	.1305	.0503	.0313	.0763	.0575
	$r = .84$		$r = .94$		$r = .96$	

pinball machine, even for worthless chips. Once again, for all three EVs, the agreement is excellent, but trends which were very slight in Figs. 1, 2, and 3 are here slightly more pronounced. In the PEV and ZEV results, the 3/8 probability of winning is slightly less liked than the 2/8. In the NEV results, the 4/8 probability of losing is slightly disliked relative to those around it (the opposite of the special preference for the 4/8 bet found in the PEV and ZEV results). These minor trends were not found

in the Harvard Ss. They are too small to make much difference in this part of the experiment, but are visible in much more pronounced form in the results of the real-money sessions.

(3) *Real money sessions.* Table III shows the results of all bets for the real-money sessions. Again there is agreement, but here the low point at 3/8 for the *PEV* and *ZEV* bets and at 4/8 for the *NEV* bets is marked. The low correlation between the *NEV* real-money results for the original and those for this replication of it should not be taken to mean that the pattern of probability-preferences is not present. The avoidance of the 4/8

TABLE III
RESULTS OF REAL-MONEY SESSIONS

Probability of winning or losing	PEV		NEV		ZEV	
	Replication	Original	Replication	Original	Replication	Original
1/8	.1347	.1116	.1177	.1957	.1518	.1399
2/8	.1356	.1265	.1250	.1719	.1607	.1592
3/8	.0909	.1376	.1688	.1652	.1144	.1563
4/8	.2175	.1830	.0706	.1116	.2045	.2009
5/8	.1282	.1406	.1331	.1049	.1315	.1339
6/8	.0747	.0900	.1672	.1265	.0820	.0677
7/8	.0950	.0975	.1315	.0871	.0552	.0804
8/8	.1234	.1131	.0860	.0372	.0998	.0618
	$r = .83$		$r = .41$		$r = .88$	

probability of winning and the preference for the 6/8 probability of winning are there, as usual. There is a tendency, though it is smaller than usual, to prefer low probabilities of losing large amounts. In the *NEV* results there is a tendency for Ss to choose in a manner opposite to their choices on the *PEV* bets. This trend was commented on in the report of the constant-*EV* experiment. It is much more marked here.

Discussion. It is evident that probability-preferences are quite reliable, in the sense that they can be reproduced from different groups of Ss under similar conditions. The peak at 4/8 and descent to 6/8 which is characteristic of the probability-preference pattern for *PEV* and *ZEV* bets shows consistently in all groups in this experiment, as does the trend towards preferring low probabilities of losing large amounts in the *NEV* bets. The general willingness to take risks increased as this experiment progressed from imagining to worthless chips to real gambling, as it did in the constant-*EV* experiment. In short, all of the major properties of probability-preferences were reproduced in this experiment.

The introduction of the valleys at 3/8 for the *PEV* and *ZEV* bets and 4/8 for the *NEV* bets, although new, is foreshadowed in the graphs of

deviations from regression lines in the constant-*EV* experiment. Presumably, these new phenomena represent intensifications of effects already present to some small degree in the old data. Why they should appear here, and in particular why they should appear in the worthless-chip and real-gambling data and not in the just-imagining data, is not clear. The most obvious hypothesis, that the pinball machine behaved in some way differently, is wrong. Every effort was made to ensure that the operation of the pinball machine was the same for this experiment as it had been for the previous experiment.

Mosteller has developed a test for use with paired-comparisons data which measures the goodness of fit between the original proportions and a set of Thurstone (Case V) scale points derived from them.⁴ This test serves as a test of the unidimensionality of the original data. Application of this test to the classroom data from this experiment results in a chi-square significant beyond the 0.01-level for each *EV*, indicating that the choices were not unidimensional, and so no Thurstone Case V scale may legitimately be derived from them. Actually, it would be most remarkable if they were unidimensional. Each bet involves two stimulus-dimensions—amount which can be won and probability of winning. It is not at all surprising that the combination of these two dimensions into one set of bets should result in non-unidimensional patterns of choices among these bets. Of course, this does not mean that the choices were made randomly. Kendall has developed a coefficient of agreement for paired comparisons which is essentially a generalization of his rank correlation coefficient (*tau*).⁵ Such coefficients were calculated for all three *EV*'s in the classroom experiment. All were significant beyond the 0.001-level, showing that *S*s definitely agreed with one another in their choices more than chance. If they agreed with one another, they cannot have chosen randomly.

EXPERIMENT II: EXTRANEOUS VARIABLES

A series of specific questions about the effects of wording, previous rolls, and other similar variables which can be expected to be important in gambling experiments was raised by the constant-*EV* experiment. A series of small experiments was designed to answer these questions. These experiments were mixed up together, and each *S* was run on all of them simultaneously. The *S*s were 11 of the 12 Harvard undergraduates used in the constant-*EV* experiment. (The twelfth had meanwhile graduated.) Each *S* ran through the deck of cards, which contained the sub-experiments described below, six times under real gambling conditions. The tapes which programmed the pinball machine were random except that on trials 30, 31,

⁴ F. Mosteller, Remarks on the method of paired comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed, *Psychometrika*, 16, 1951, 207-218.

⁵ M. G. Kendall, *Rank Correlation Methods*, 1948.

and 32 all Ss rolled 4; on trial 33 they rolled 1 or 6; and on the last trial they rolled 2 or 8. Three different tapes with these properties were prepared, and each was used on two of the six days.

Wording of Zev bets. One of the major findings of the constant-*EV* experiment was that the *ZEV* results very much resembled the *PEV* results. (This finding was, of course, confirmed in the replication reported above.) The obvious hypothesis to explain this finding, based on the fact that the *ZEV* bets read like the *PEV* bets except that possibilities of losing are added, is that Ss in reading a bet pay attention for the most part only to the statements about what they can win, rather than those about what they can lose. In the constant-*EV* experiment, however, the *ZEV* bets were always written with the statement about winning first and the statement about losing following, e.g. "If you roll a 2, a 3, a 5, a 7, or an 8, you win \$0.84. If you roll anything else, you lose \$1.40." Perhaps Ss pay attention to whatever comes first, regardless of whether it is the winning or the losing part of the bet. To check

TABLE IV
RESULTS OF ZEV REVERSED WORDING EXPERIMENT

Probability of winning	Reversed wording	Normal wording
1/8	.1677	.1399
2/8	.1423	.1592
3/8	.1412	.1563
4/8	.1766	.2009
5/8	.1033	.1339
6/8	.0909	.0677
7/8	.1001	.0804
8/8	.0779	.0618

$$r = .89$$

this possibility, a set of *ZEV* cards was prepared for this experiment in which the order was reversed; the possibility of loss was named first and the possibility of winning came afterwards.

Results. Table IV shows the results of Experiment II. The real-gambling *ZEV* results of the original experiment are included for comparison. It is clear that the results of reversed wording are very similar to the results of the previous wording, so it is reasonable to conclude that the order of wording does not make much difference. It seems likely that Ss simply pay more attention to statements about what they can win than to statements about what they can lose, regardless of wording or location, at least so far as properties of wording and location are changed by this experiment.

Payoff cells and number of bets. The eight cells at the base of the pinball machine, into one of which the ball must roll, are numbered 1 through 8 from left to right. It seems quite likely that preferences for position and number are involved in Ss' choices. Any exhaustive test of this possibility would require a tremendous amount of experimentation, but a miniature exploration of it was attempted.

Bets were selected for the 2/8, 4/8, 6/8, and 8/8 probabilities of winning or losing. Table V shows these bets. The cells which pay off are very different from those which paid off in the corresponding bets in the constant-*EV* experiment. There are, of course, only four bets, and therefore six pairs, at each *EV*.

The results are shown in Fig. 4. Two curves representing the results of the constant-*EV* experiment are included in each section of the figure for comparison. The one called 'transformed' is made up of the points corresponding to these probabilities on the appropriate curves in the constant-*EV*

TABLE V
WORDING OF BETS IN CHANGED PAYOFF CELLS EXPERIMENT

Positive Expected Value

- (1) If you roll a 1 or an 8, you win \$2.10. If you roll anything else, you win nothing.
- (2) If you roll a 1, a 2, a 7, or an 8, you win \$1.05. If you roll anything else, you win nothing.
- (3) If you roll anything but a 4 or a 5, you win \$0.70. If you roll a 4 or a 5, you win nothing.
- (4) Regardless of what you roll, you win \$0.53.

Negative Expected Value

The four bets with negative expected value are identical with those listed above except that the verb in each bet is "lose" instead of "win."

Zero Expected Value

- (1) If you roll a 1 or an 8, you win \$2.10. If you roll anything else, you lose \$0.70.
- (2) If you roll a 1, a 2, a 7, or an 8, you win \$1.05. If you roll anything else, you lose \$1.05.
- (3) If you roll anything but a 4 or a 5, you win \$0.70. If you roll a 4 or a 5, you lose \$2.10.
- (4) Regardless of what you roll, you neither win nor lose.

experiment. The curve called 'calculated' is based on the fact that each pair of bets used in this part of this experiment is identical in probability and amount of payoff with one of the pairs used in the constant-*EV* experiment. The corresponding pairs in the constant-*EV* experiment were re-scored as though they had been a replication of this experiment, and the result is the calculated points. It is readily apparent that in all cases agreement between this and the constant-*EV* experiment is excellent, in spite of differences in wording and in the number of bets being paired with one another.

A fourth set of bets was also used, just like these *ZEV* bets except that the wording was reversed, as in the previous section of this experiment. This procedure was intended as insurance against the possibility that the previous experiment might show systematic differences due to wording. The results of the fourth group were indistinguishable from the *ZEV* results for this experiment presented in Fig. 4, giving further weight to the conclusion about the effect of reversed wording on *ZEV* bets which was presented above.

In general, so far as the limited procedures of this experiment go, it is

possible to conclude that number-preferences, position-preferences, and the number of alternative bets are not important determiners of choices in situations like that used here and in the previous experiments.

Effect of previous roll. It seems likely that *Ss* are influenced in their choices by the number rolled on the pinball machine on preceding trials. In reports after sessions in previous experiments, *Ss* have named this as one of the determiners of

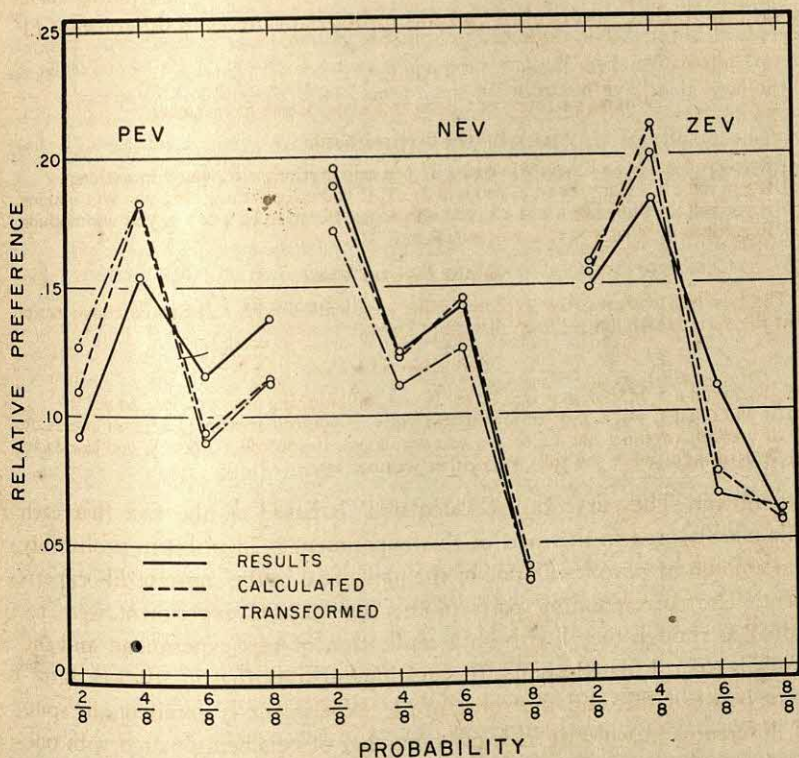


FIG. 4. RESULTS OF CHANGED PAYOFF-CELL EXPERIMENT

(The curves labelled "Results" show the results of this experiment. The curves labelled "Transformed" are taken from the constant-*EV* experiment. The curves labelled "Calculated" are calculated from the constant-*EV* experiment.)

their choices, but observation of their behavior raised some question as to whether they were correct in their convictions. To test this theory thoroughly would require a tremendous amount of experimental work, but an approximate test for one special case was easy to make. Two identical cards were prepared. Alternative *A* on each case was "If you roll a 4, you win \$4.20. If you roll anything else, you win nothing." Alternative *B* was "If you roll a 2, a 3, a 5, a 7, or an 8, you win \$0.84. If you roll anything else, you win nothing." Alternative *A* was, of course, selected because

it is the only bet in the usual list of *PEV* bets which involves only one winning cell on the pinball machine. Alternative *B* was a neutral alternative, which did not pay off on 4.

Of these two duplicate cards, one was presented to *Ss* as the first card in the deck. The other was presented as the thirty-third card. On both cards, the tape of the pinball machine programmer was arranged to give a 1 or a 6, so that no question of reinforcement could arise. On the thirtieth, thirty-first, and thirty-second cards, every *S* rolled a 4. The cards for those places in the series were so chosen that a 4 neither won nor lost any money (reinforcement could not occur). It might be supposed, then, that if the previous rolls are important, some of the *Ss* who preferred alternative *A* on the first card might switch to *B* on the second card, on the hypothesis that "I've just rolled 4 three times; I can't possibly do it again."

The results of this part of the experiment are given in Table VI. It is apparent that more people switched away from the long shot (9) than

TABLE VI
CHOICES OF ALTERNATIVES A AND B IN THE EXPERIMENT ON PREVIOUS ROLLS

First roll	Second roll	
	A	B
A	18	2
B	9	37

switched toward it (2). A test of whether the off-diagonal terms are equal, computed with the exact binomial, comes out not quite significant ($p = 0.065$).

It seems likely that the tendency to change in favor of the long shot on the second roll was caused by the excitement of the ongoing gambling, which encouraged *Ss* to be less cautious than they had intended to be. The general conclusion to be drawn from this part of the experiment is that so far as this very special and limited test is concerned, the previous roll does not have too much effect on choices in subsequent rolls. This conclusion was borne out by the spontaneous comments of two *Ss*, who said when choosing on card 33, "I can't possibly roll another 4—but I'll take *A* anyhow. It's only money."

Extreme bet. One approach to the problem of the effect of changing the amount of money involved in a bet is to compare bets of the same *EV* but with even greater extremes of probability of winning than those nearest them in the range of already-familiar bets. This was done in a very tentative way. One card was prepared. Alternative *A* was "If you roll a 4, you win \$4.20. If you roll anything else, you win nothing." Alternative *B* was "If you roll a 1, a 3, a 5, or a 7 on the first roll, you take a second roll. If you roll a 4 on the second roll, you win \$8.40. If you are not successful on both rolls, you win nothing." This card was the very last card

in the deck. It was put there because some *Ss* were so far behind that it was desirable to cause them to win \$8.40, and so they were allowed to win it on the last day, when it could not affect their subsequent behavior. Out of 66 choices made by all *Ss* on this card, 64 were for alternative *B*. The result seems due to the fact that this was the last bet. *Ss* did not expect to win on it, according to their comments, and in fact until the last day no one did win on it, since all *Ss* rolled a 2 or an 8 on the first roll; nevertheless, all *Ss* liked this card, looked forward to it, and expressed glee at the chance of winning \$8.40 when it turned up. It would have been interesting to try a similar card with a negative instead of a positive *EV*, but any subsequent experimentation of this sort should not put the card at the end of the series of bets.

General conclusion. In Experiment II, no evidence was found to destroy confidence in probability-preferences as the explanation for the results obtained in the constant-*EV* experiment and Experiment I. Wording, choice of numbers to pay off, and previous rolls proved to be (in the sketchy tests given them by this experiment) relatively unimportant in determining gambling decisions.

EXPERIMENT III: NON-GAMBLING RISKS

All experiments so far in this series were done with a pinball machine, poker chips, and monetary bets of a particular kind—in short, all of them have been quite similar. The repeated discovery of the same pattern of preference in similar situations shows that the pattern is recoverable from different *Ss* and under small variations in experimental conditions. It is desirable, however, to find out in how many places in the field of risk-taking behavior this pattern may be expected to occur, particularly whether it will occur in a risky non-gambling situation. To study this question, an experiment was designed which transferred all the formal properties of the constant-*EV* experiment (and its replication here) to a totally different, non-gambling situation.

Apparatus. A visual display was the principal apparatus. It consisted of a painting on glass of a map of an island. On this island's shores, spaced at approximately equal intervals and located on indentations of the shore-line, were eight ports, each with a neutral-sounding two syllable English name (e.g. Cumnock, Hartfell, Gosforth, and the like). On a table in front of *S* were eight buttons, each with the name of a port on it; a chute for delivering 'chips'; a hole for returning them (the chute and hole were connected with a conveyor belt, which carried the chips to and from *S*); and a counter. When *S* pressed a button, the area around the corresponding port on the map lighted up. *E* could turn on a much more intense light in the center of the port itself, and could extinguish all lights and at the same time operate the counter.

Subjects. Two groups of *Ss* were used in this experiment. One was the same

group of 11 randomly chosen Hopkins undergraduates described in Experiment I of this paper. Experiment III was conducted after Experiment I and several other studies. Another group of 12 randomly chosen Johns Hopkins undergraduate men ran through this experiment first, and afterwards through two sessions of gambling for worthless chips on the pinball machine under conditions exactly like those of Experiment I.

Procedure. Each *S* in this part of the experiment was told:

"You are an admiral in the U. S. Navy, assigned to command a small amphibious task force. The mission of your task force is to capture as much as possible of the output of an enemy ammunition factory located in the interior of this island. You do not have the resources to attack the factory directly, but you know that its output must be shipped from one of the island's eight ports. Consequently, your general strategy is to attack one or more of these ports in hope of capturing this ammunition.

"Your orders from your superior officers will be presented to you in a form which leaves you two alternatives, one of which *has* to be executed and the consequences accepted. You will be ordered to use your forces in one of three types of attacks on these ports.

(1) *Sneak attack.* You attack one or more ports with small numbers of men, quietly, and with no preliminary bombardment, hoping to surprise the enemy and capture the ammunition shipment. If you attack the port where the ammunition shipment is, then you succeed. Otherwise you fail, but lose no ammunition.

(2) *Preventive attack.* You will at times be ordered to attack when the ammunition shipment is too well guarded to capture. In such cases, you have of course no hope of gaining ammunition. If, however, you attack the port where the ammunition is located, then you will fight with the guard, and consequently lose ammunition. Therefore your purpose in this kind of attack is to attack ports where the ammunition shipment is *not* located, if possible. In this type of attack, you can lose ammunition, but cannot gain any.

(3) *Attack in force.* You make a full-dress attack, with intensive preliminary bombardment. If you do not attack the port where the ammunition shipment is located, then you lose ammunition, because you have used it in the attack. If you do attack the port where the shipment is located, then you gain substantial amounts of ammunition over and above the amount used in the attack."

Of course, each sneak attack corresponds to a *PEV* bet, each preventive attack corresponds to a *NEV* bet, and each attack in force corresponds to a *ZEV* bet. Table VII shows the actual wording of three of the twenty-four different attacks. From here on all these attacks will be referred to as though they were *PEV*, *NEV*, or *ZEV* bets.

Each bet was paired with all others of its *EV* according to Ross's optimal order for pairs in the method of paired comparisons.⁶ Each pair was typed on a card, and the resulting deck of cards was given to *S*, along with chips to represent 200,000 pounds of ammunition. Three different kinds of empty cartridge cases were used as the chips, each kind representing a different amount of ammunition. If *S* won a bet, the proper number of chips was delivered to him by the conveyor belt; if he lost, he inserted the proper number of chips into the hole and the conveyor belt carried it through the wall to *E*. When *S* decided which of the two alternatives of a card he preferred, he pressed the buttons which had on them the names of the ports which that alternative required him to attack. *E*, in the next room, then turned on a more intense light which indicated in which port the ammunition was actually

⁶ *Op. cit.*

located, and *S* received or gave up chips accordingly. Then *E* cleared the display and operated the counter, and *S* went on to the next card.

Results. There are no important differences between the group which was run on the pinball machine first and the island display second and the group which was run in the reverse order, so the data from the two groups are pooled here. The results from the *PEV*, *NEV*, and *ZEV* bets

TABLE VII

SAMPLES OF THE ATTACKS USED IN THE ISLAND DISPLAY EXPERIMENT

Corresponds to PEV Bet

- (1) This is a sneak attack. You attack Hartfell, Duffus, and Frinton. If the ammunition shipment is in one of the attacked towns, you capture 14,000 pounds of ammunition. If it is not, you lose no ammunition.

**Corresponds to NEV Bet*

- (2) This is a preventive attack. You attack every town except Nayland. If the ammunition shipment is in one of the attacked towns, your forces are repelled and you lose 6,000 pounds of ammunition. If it is not, you lose no ammunition.

Corresponds to ZEV Bet

- (3) This is an attack in force. You attack Cumnock and Gosforth. If the ammunition shipment is in one of the attacked towns, you capture 21,000 pounds of ammunition. If it is not, you expend 7,000 pounds of ammunition.

on the island display compared with the corresponding bets on the pinball machine are shown in Figs. 5, 6, and 7, respectively. It is at once apparent that the tremendous differences between the two situations have produced substantial differences in patterns of choices, but these differences are not so great as to obliterate the main features of the probability-preference pattern. In Fig. 5, the relative peak at $4/8$ and the relative valley at $6/8$, which are the main features of *PEV* probability-preferences, can still be found in the island-game results. In Fig. 6, the trend to prefer bets with low probabilities to those with high probabilities, which is characteristic of *NEV* probability-preferences, appears in the island-game results, although the $1/8$ point violates it. In Fig. 7, where the agreement is least good, the island-game results still show a very small relative peak at $4/8$ and a similarity of trend.

It is interesting to note that the general rule found in the constant-*EV* experiment; namely, that the *PEV* and *ZEV* results are very much alike, holds here also. Indeed, the *PEV* and *ZEV* results with the island game are almost as much like each other as they are like the pinball results with which they are compared.

In spite of the differences between the island game and the gambling situations in which probability-preferences were discovered, elements of

the probability-preference pattern can still be found in the island-game results. It seems likely that those elements which did show up in the island-game results can be regarded as general features of responses to risks, not confined to gambling with the pinball machine, or indeed to gambling itself.

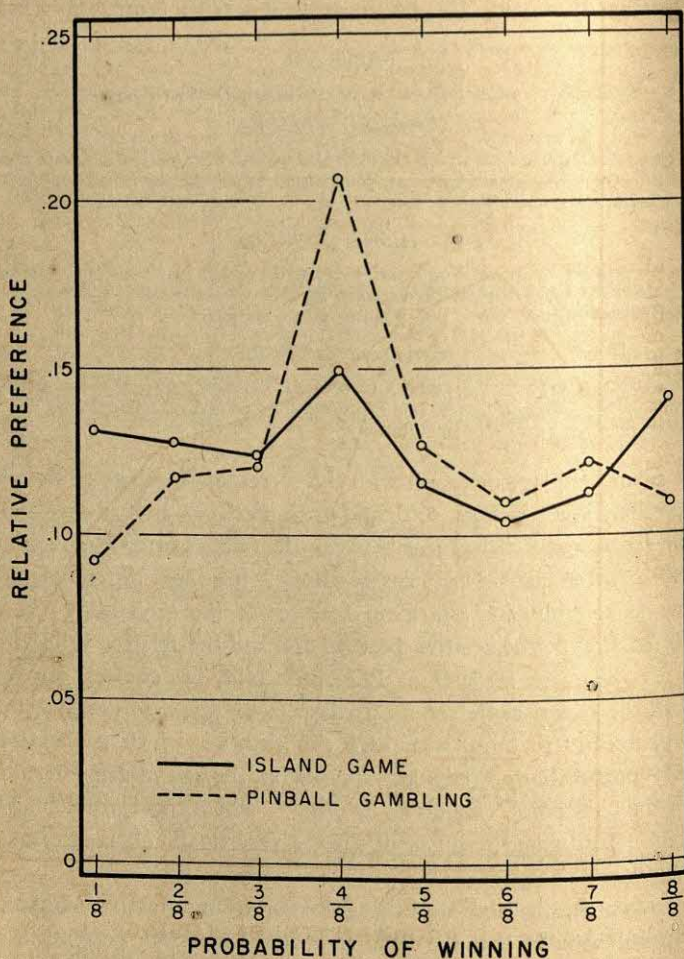


FIG. 5. *PEV* ISLAND-DISPLAY RESULTS COMPARED WITH *PEV* PINBALL RESULTS

EXPERIMENT IV: DIFFERENT-*EV* LEVELS

All of the experiments described so far in this paper, and all of those described in the report of the constant-*EV* experiment, were based on an

EV of \$0.525 per roll. The probability of winning or losing and the amount of money which could be won or lost have been tied variables, since whenever one is changed the other also must be changed in order to keep the *EV* constant. Theoretically, therefore, all of the results found so far,

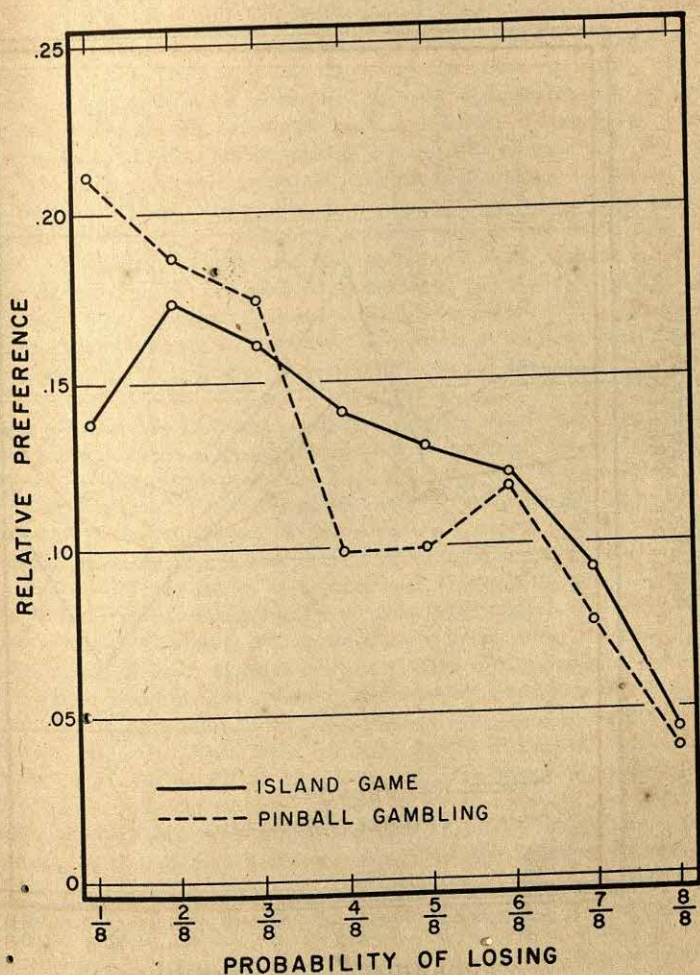


FIG. 6. *NEV* ISLAND-DISPLAY RESULTS COMPARED WITH *NEV* PINBALL RESULTS

which have been attributed to preferences among probabilities, might instead have been attributed to preferences among the amounts of money involved in the bets. The report of the original constant-*EV* experiment discusses in more detail how the concept of the subjective value of money

(which economists call the utility of money) could be used to account for these results. The crux of this discussion is that for the results so far it would be possible to draw an interval scale of the utility of money which would predict the choices which Ss actually made in these experiments.

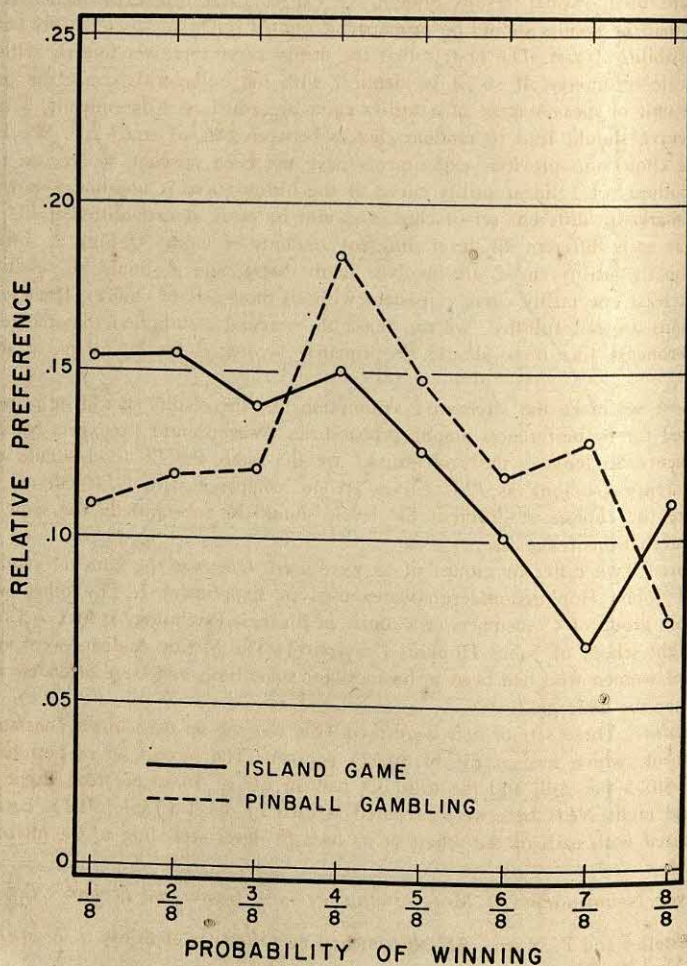


FIG. 7. ZEV ISLAND-DISPLAY RESULTS COMPARED WITH ZEV PINBALL RESULTS

Such a scale would have several inflection points and so would look a bit peculiar, but it could be drawn. A method for determining these scales experimentally has been proposed in very general form by von Neumann

and Morgenstern⁷ and put into practice by Mosteller and Nogee.⁸ The difficulties for this method raised by the phenomenon of probability preferences have been discussed in the previous papers.

In this experiment, bets located at several different positive and negative *EV* levels are used.⁹ What results should we expect from this experiment? If we assume that the results should be primarily accounted for by means of utility curves, two possibilities exist. The first is that the utility curve increases linearly with the dollar-scale of money. If so, it is identical with the dollar-scale, since the origin and the unit of measurement of a utility curve are arbitrary. Consequently, a linear utility curve should lead to random choices between bets of equal *EV*. We know that the choices in previous experiments have not been random, so we can reject the hypothesis of a linear utility curve. If the utility curve is non-linear, however, then a markedly different set of choices should be made at each different *EV*-level (since at each different *EV*-level different amounts of money, falling at different places on the utility curve, are involved in the bets), and it should be possible to draw at least one utility curve consistent with all these sets of choices. If these two predictions are not fulfilled, we can reject the original assumption that the results of experiments like these should be primarily accounted for by means of utility curves.

Suppose we make the alternative assumption that the results should be primarily accounted for by preferences among probabilities. What should we expect from this experiment? So long as the probabilities are the same, the *EV*-level should make no difference (so long as *PEV* curves are not compared with *NEV* curves), and therefore the choices at different *EV* levels should be substantially the same. We shall find that this is the case.

Subjects. Two different groups of *Ss* were used. One was the same 11 randomly selected Johns Hopkins undergraduates used in Experiment I. The other was a classroom group of 43 members of a course in Business Psychology at McCoy College (the night school of Johns Hopkins University). The McCoy students were mostly men and women who had been in business for some time, and were of course much older than the undergraduates.

Procedure. Three sets of bets were used. One was the set used in the constant-*EV* experiment, which has an *EV* of \$0.525 per roll. The second set of bets had an *EV* of \$0.75 per roll, and the third set had an *EV* of \$0.30 per roll. Eight *PEV* bets and eight *NEV* bets were prepared at each *EV*-level (Table VIII). Each bet was paired with each of the others at its own *EV*-level according to the method of

⁷ J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, 1944.

⁸ Mosteller and P. Nogee, An experimental measurement of utility, *J. Polit. Econ.*, 59, 1951, 371-404.

⁹ In this experiment, each choice was between the members of a pair of bets each of which had the same *EV* as the other. In one of the previously reported experiments, each choice was between the members of a pair of bets in which one bet had a larger *EV* than the other. These two different experimental arrangements of course give very different results. For the experiment in which the members of each pair of bets had different *EV*s, see Edwards, Probability preferences among bets with differing expected values, *op. cit.*, 56-67.

paired comparisons. Thus, there were six different sets of 28 pairs of bets each, having the *EVs* +\$0.75, +\$0.525, +\$0.30, -\$0.30, -\$0.525, and -\$0.75 per roll.

The classroom group was shown a slide of the pinball machine, told about its functioning (but not that it was rigged), and then shown slides of each of the 168 pairs of bets and asked to imagine they were gambling. This procedure is identical with that used in the classroom experiment described in Experiment I. Of course the classroom *Ss* made their judgments only once. The individually run Hopkins

TABLE VIII

BETS USED IN EXPERIMENT IV

Positive Expected Value = \$0.75

- (1) If you roll a 4, you win \$6.00. If you roll anything else, you win nothing.
- (2) If you roll a 1 or a 7, you win \$3.00. If you roll anything else, you win nothing.
- (3) If you roll a 2, a 4, or a 6, you win \$2.00. If you roll anything else, you win nothing.
- (4) If you roll a 2, a 4, a 7, or an 8, you win \$1.50. If you roll anything else, you win nothing.
- (5) If you roll a 2, a 3, a 5, a 7, or an 8, you win \$1.20. If you roll anything else, you win nothing.
- (6) If you roll anything but a 3 or a 6, you win \$1.00. If you roll a 3 or a 6, you win nothing.
- (7) If you roll anything but a 5, you win \$0.86. If you roll a 5, you win nothing.
- (8) Regardless of what you roll, you win \$0.75.

Negative Expected Value = \$0.75

The eight bets with a negative expected value of \$0.75 are identical with those listed above except that the verb in each bet is "lose" instead of "win."

Expected Values of \$0.525 and \$0.30

The bets at these levels of expected value are identical with the two sets of bets described above except for the amounts of money involved. The amount of money involved in each bet at each *EV* level was as follows:

Probability	\$0.75	\$0.525	\$0.30
1/8	\$6.00	\$4.20	\$2.40
2/8	3.00	2.10	1.20
3/8	2.00	1.40	0.80
4/8	1.50	1.05	0.60
5/8	1.20	0.84	0.48
6/8	1.00	0.70	0.40
7/8	0.86	0.60	0.34
8/8	0.75	0.53	0.30

undergraduates really gambled on the pinball machine in the manner described in Experiment I. Each session consisted of one run-through of all pairs of bets with positive or negative *EVs* of \$0.75 and \$0.30, and there were four different sessions for each *S*. (The \$0.525 *EV* bets were not included because these same *Ss* had already made four judgments of each pair of them in Experiment I; the data on the \$0.525 *EV* bets shown in Figs. 8 and 9 come from Experiment I.)

Results. Fig. 8 shows the *PEV* results for the individually run *Ss*. The usual features of probability-preferences are all present. There is a peak at 4/8, a valley at 6/8, and the valley at 3/8 which has shown up in all real gambling experiments using Hopkins *Ss*, although it did not show up in

the original experiment using Harvard Ss. The only major difference is that the \$.75 EV has its valley at 7/8 instead of 6/8. It is obvious that the three curves are very much alike. Fig. 9 shows the NEV results for the individually run Ss. Again the three curves are very much alike. As

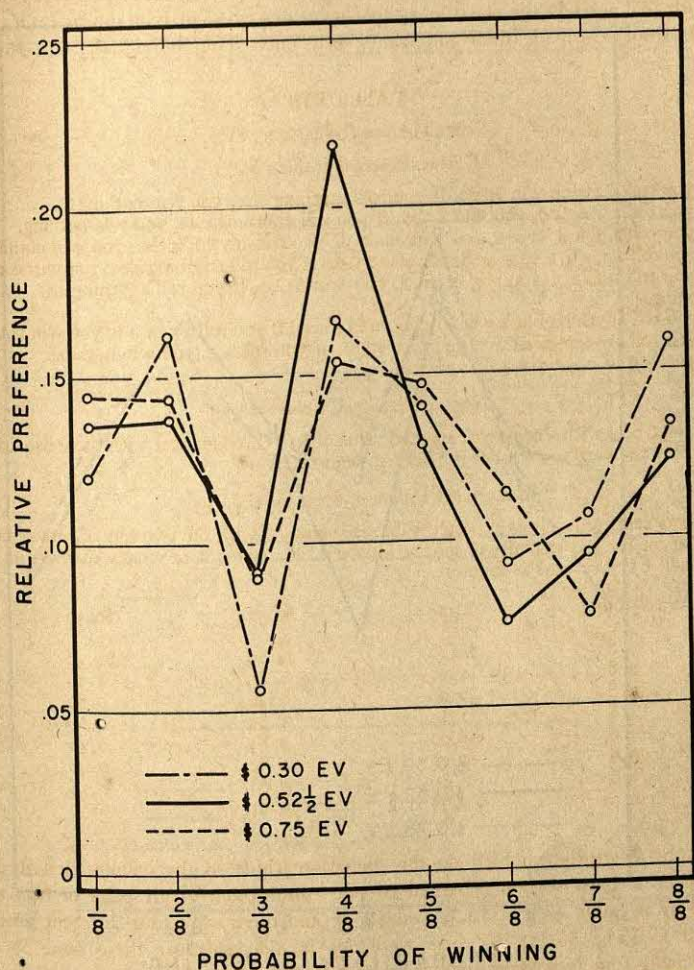


FIG. 8. PEV RESULTS FOR INDIVIDUALLY RUN Ss AT THREE DIFFERENT EV-LEVELS

usual for real gambling with Hopkins Ss, the NEV curves look like the inverse of the PEV curves.

In Figs. 10 and 11 the PEV and NEV results for the classroom group are shown. Although much reduced, the usual peak at 4/8 and valley at

6/8 are present in the *PEV* curves. The \$.75 curve is again slightly deviant. The *NEV* curves are very much like all previous *NEV* curves. The differences between the results for the individually run *Ss* and the classroom group, of course, should be attributed to the fact that the members of the

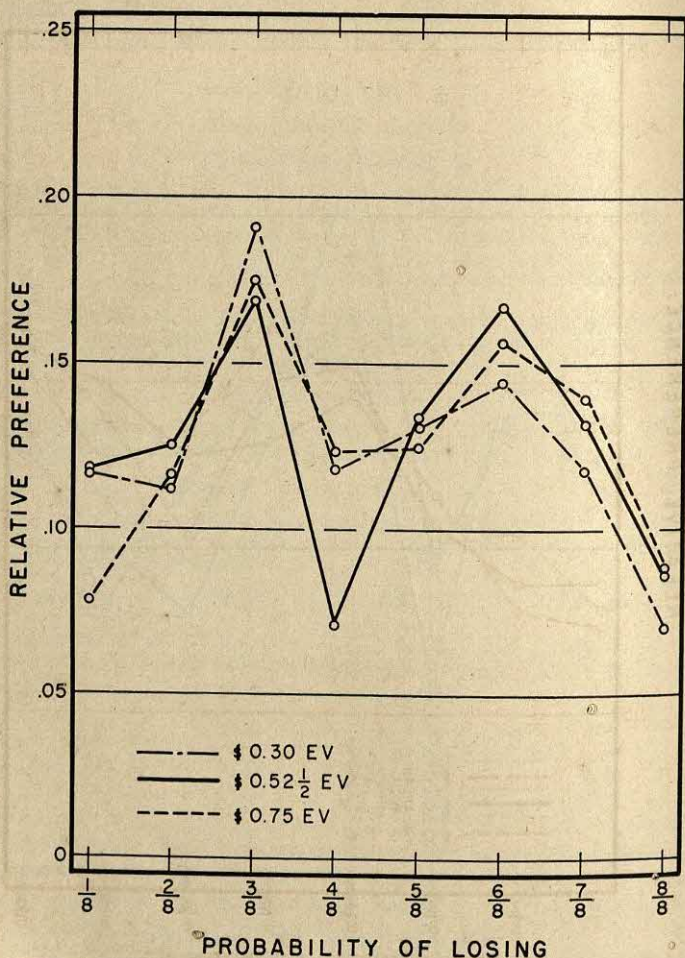


FIG. 9. *NEV* RESULTS FOR INDIVIDUALLY RUN *Ss* AT THREE DIFFERENT *EV*-LEVELS

classroom group were just imagining that they were gambling, while the individually run *Ss* were gambling for real money. Similar differences were found between these two conditions in the constant-*EV* experiment. Table IX shows the correlations among the various sets of results.

Discussion. Figs. 8 and 9 compare sixteen triplets of bets. Each member of a triplet has the same probability of winning as the other two members, but the amount of money involved in it and its *EV* are different from the other two. Nevertheless, all three members of a triplet were preferred

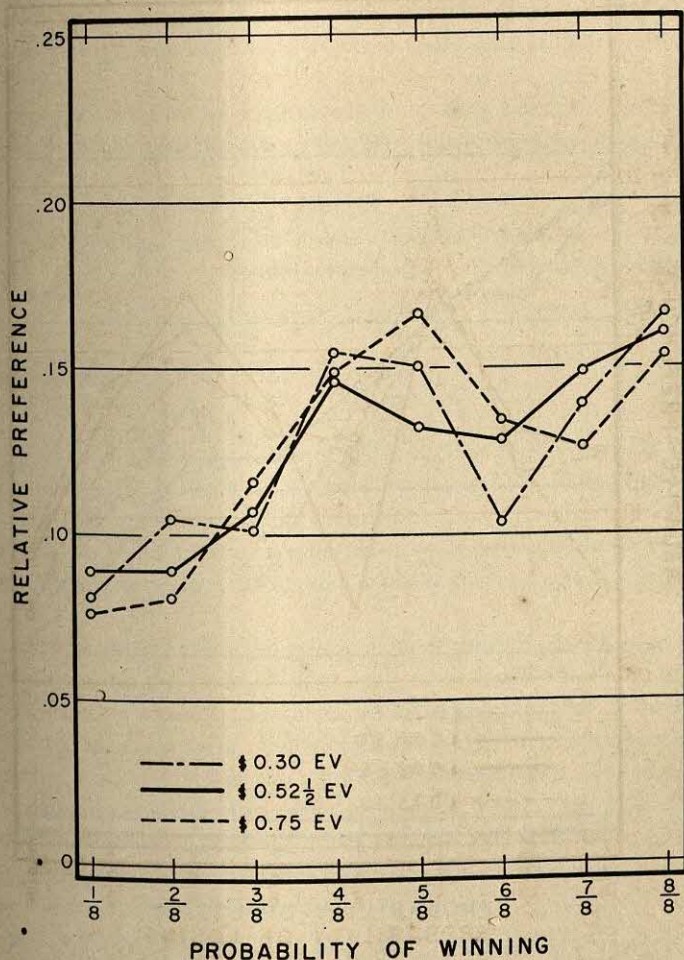


FIG. 10. PEV RESULTS FOR CLASSROOM Ss AT THREE DIFFERENT EV-LEVELS

to about equal extents to the other bets of the same *EV* with which they were compared. This was true for all sixteen triplets. The inevitable conclusion is that these preferences were based on the probabilities involved in

the bets, which the members of a triplet have in common, rather than the amounts of money involved, which are different for each member of a triplet.

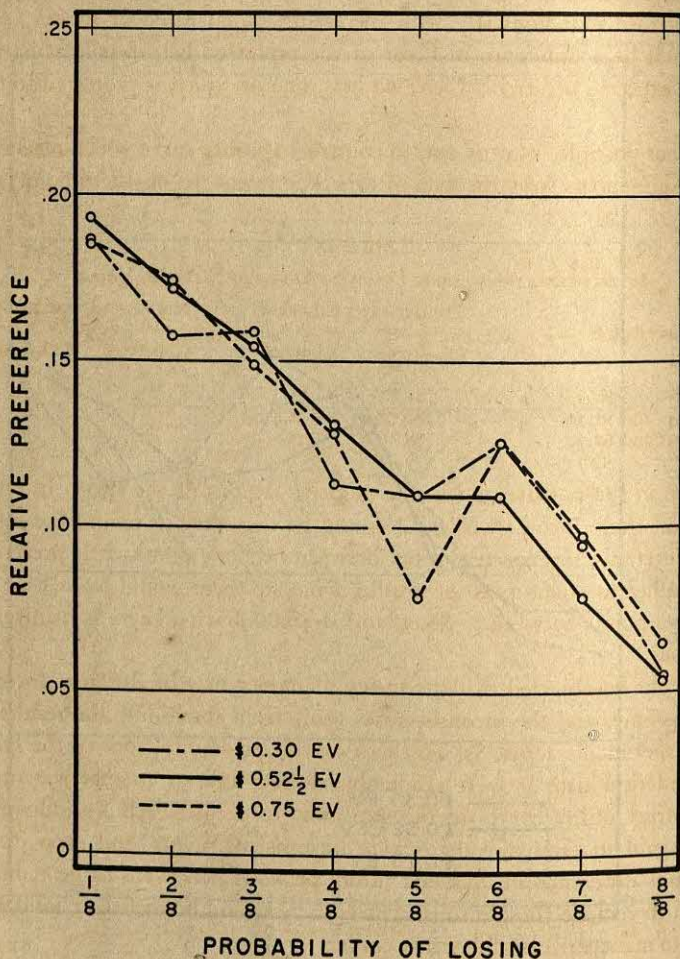


FIG. 11. *NEV* RESULTS FOR CLASSROOM Ss AT THREE DIFFERENT *EV*-LEVELS

These results rule out all possibility of explaining the phenomenon which has been called 'probability-preferences' throughout this series of papers by appeal to the utility of money. An attempt was made to construct a group utility curve which would account for the data from this

experiment under the most permissive reasonable rules. These rules were: the utility of \$0.00 is 0 utiles, the utility of +\$0.30 is 30 utiles, the utility of -\$0.30 is -30 utiles (these assumptions determine scale and origin, but do not otherwise affect the shape of the curve), the utility of money must increase monotonically with increasing dollar-value of money, and there must be a difference in favor of the preferred bet of at least half a utile in expected utility between two bets, one of which is preferred to the other.¹⁰

It is not possible, it turns out, to construct a utility curve which meets all these requirements from the data of this experiment. Inconsistent points oc-

TABLE IX
INTERCORRELATIONS AMONG DIFFERENT LEVELS OF EXPECTED VALUE

Correlation	Positive expected value		Negative expected value	
	Classroom	Individual	Classroom	Individual
\$0.75 and \$0.525	.85	.72	.94	.70
\$0.75 and \$0.30	.83	.78	.93	.82
\$0.525 and \$0.30	.88	.73	.97	.77

cur both in the positive and in the negative regions of the curve. In order to correct these inconsistencies, it would be necessary to reduce the minimum difference in expected utility between bets one of which is preferred to the other to about 1/20 of a utile. Even so, there would be at least 11 inflection points between -\$6.00 and +\$6.00. Such a curve is manifestly ridiculous.

It might be objected that the utility of money may be different for different people, and the inconsistencies come from combining the results of these experiments across Ss. To check on this possibility, four of the Hopkins undergraduate Ss were randomly selected, and an attempt was made to construct utility curves in the same way to predict each S's individual data. Again in each case there were inconsistent points, and even if the minimum difference in expected utility between bets were to be reduced enough to remove these inconsistencies there would be too many inflection points to make a reasonable utility curve.

It must be concluded that these data cannot be accounted for by appeal

¹⁰ "Utile" is the convenient name used by Mosteller and Nogee for their arbitrary unit of utility. The utiles discussed here, however, are smaller than those of Mosteller and Nogee, since Mosteller and Nogee defined -1 utile as the value of -\$0.05, while I am defining 30 utiles as the value of +\$0.30 and -30 utiles as the value of -\$0.30.

to the notion of the utility of money. Preferences among probabilities, then, seem to be the only reasonable explanation.

It would not be safe to assume that differences between subjective and objective values of money played no rôle in the results. There are, after all, differences between the three different expected values. These differences are relatively small. It seems quite reasonable, however, to assume that they should be attributed to the operation of utilities. Verification of this assumption must wait on the development of a satisfactory method for measuring utilities. Any such method, if it uses risk-situations, will have either to measure and allow for probability-preferences (in which case a separate method of measuring probability-preferences must first be developed, one in which utility plays no role) or else will have to hold probabilities constant. The latter seems the more promising possibility.

GENERAL CONCLUSIONS

The phenomenon of probability-preferences has been shown to occur in a wide variety of experimental results. The essential condition for its occurrence seems to be the presence of a risk-situation in which there is a well-defined probability of gaining or losing something. It shows much the same characteristics regardless of the experimental conditions. Whether other kinds of risks, notably sequential risks (such as a rat takes in a series of trials in a T-maze), will also show the phenomenon of probability-preferences is a question which can, of course, only be answered by further research.

SUMMARY

A previous experiment has shown that choices between bets of equal expected value stated in terms of rolls on a pinball machine are influenced by preferences among the probabilities involved. This and subsequent research raised several methodological questions, which the four experiments here reported were designed to examine.

Experiment I showed that a replication of the previous experiment with different Ss and slightly different experimental conditions gave essentially the same results. The pattern of preferences among probabilities reported in the previous experiment was found. The choices involved could not be fitted to a unidimensional subjective scale. The Ss agreed with one another to a very highly significant extent.

Experiment II showed that choices between pairs of bets were not substantially influenced by the wording of the bets, the outcome of pre-

vious rolls of the pinball machine, or the number of bets being compared.

Experiment III showed that the essentials of the pattern of preferences among probabilities could be found in a non-gambling game involving risk, designed to be as different as possible from the pinball situation.

Experiment IV showed that the same pattern of preferences among probabilities could be recovered at several different levels of expected value. It was not possible to find any reasonable scale of the subjective value of money which would be consistent with the results of the experiment. This finding was interpreted to mean that the results of experiments like these should be attributed primarily to preferences among probabilities rather than to differences between the subjective and the objective value of money.

COMPARABILITY OF THE METHOD OF SINGLE STIMULI AND THE METHOD OF PAIRED COMPARISONS

By SIDNEY M. NEWHALL, Eastman Kodak Company

This study is a comparison of two preferential judging methods for the purpose of estimating the interchangeability of the methods as used by appropriately instructed observer groups. The method of paired comparisons commonly consists in juxtaposed comparisons of given related items in all possible pairs whereas the method of single stimuli consists in simple successive sorting of the given items into several prescribed categories. Since the method of rating, traditionally applied for complex and esthetic evaluations, and the psychophysical method of so-called absolute judgment are now recognized to be essentially the same, the expression, *method of single stimuli*, is here used as a general term covering both.

Kodacolor prints and projected Kodachrome transparencies were the items judged in this study of the most widely used visual judging methods in the photographic industry. When color-print inspectors sort photographs one at a time into such categories as 'accept' and 'reject,' they are using the method of single stimuli. When experts attempt to determine the relative quality of several photographs by directly comparing each with every other they are using the method of paired comparisons.

The method of paired comparisons has long had the reputation of superiority in sensitivity and precision, but it is by far the more laborious and time-consuming when more than a very few items are involved. For instance, the full method of paired comparisons, with the space error compensated, requires twice as many observations as the method of single stimuli when there are as few as 3 items to be judged; 4 times as many observations when there are 5 items; 9 times as many observations when there are 10 items; and so on. There are various advantages of the method of single stimuli, in addition to brevity, and these have been summarized by Guilford.¹ There is, then, a significant efficiency problem in determining whether the two methods yield sufficiently comparable results in actual judging situations with the materials of interest to justify more extensive use of the one method at the expense of the other.

* Accepted for publication December 8, 1952.

¹ J. P. Guilford, *Psychometric Methods*, 1936, 279-280.

There is considerable reason to suspect that a close relation might exist. First, the methods themselves are more similar judgmentally than would at first appear. In both methods the individual develops a reference background which affects his judgments. In case of the method of single stimuli this reference background is sometimes perfected deliberately by preliminary observations, but in any case the background will develop during the course of the experimentation. Secondly, in both methods the various items are likely to be observed in random order and each judgment may be most influenced by the most contemporaneous items. The chief difference will be that these items are present simultaneously in the one method and in close succession in the other.

Three experimental studies have been reported which bear directly upon the point at issue. The earliest is that of Valentine on the pleasingness of musical intervals.² The rank difference correlation, ρ , of his results with 5 *O*s and 12 intervals was 0.974. In the method of paired comparisons, however, his *O*s were given the unusual instruction "to judge each interval upon its own merits, and to avoid as far as possible being influenced by the interval with which it was compared." Such an instruction tends to get away from the intent of the method of paired comparisons and to make the judging abnormally like that of the method of single stimuli. Such an instruction might have produced atypically close agreement of results.

The methods study by Tinker included comparison of the graphic rating method, which is a form of the method of single stimuli, with the method of paired comparisons.³ In the first part of this study 30 *O*s judged 10 color patches with respect to affective value; and the group-results for the two methods correlated 0.835. In the second part groups of 16 and 20 *O*s judged musical intervals by both methods; and the correlations were 0.917 and 0.861. Because of interest in the true relation between the methods, Tinker also reported his correlations as corrected for attenuation. These corrected coefficients were, respectively 0.973, 0.956 and 0.932.

The Barnhart study employed 41 *O*s with two series of items, viz., 10 black and white drawings and 10 geometrical forms.⁴ In this study the preferential frequencies had been converted to probability scale values prior to correlation. The Pearson correlations between the methods of single stimuli and paired comparisons in the two series turned out to be 0.968 and 0.954, respectively.

PART I

Our general procedure was to secure from a number of *O*s comparable preferential judgments by the methods of single stimuli and paired comparisons on several series of Kodacolor Prints of varied quality. Then the degree of correlation of the results of the two methods was computed to

²C. W. Valentine, The method of comparison in experiments with musical intervals and the effect of practice on the appreciation of discords, *Brit. J. Psychol.*, 7, 1914, 118-135.

³M. A. Tinker, Methods employed for ranking stimuli, *J. Gen. Psychol.*, 5, 1931, 53-62.

⁴E. N. Barnhart, A comparison of scaling methods for effective judgments, *Psychol. Rev.*, 43, 1936, 387-395.

determine whether the results were sufficiently similar to permit the substitution of methods.⁵

Choice of materials. Three series of 11 prints each were assembled from various sources for the experimental comparison of methods. A limited attempt was made to secure a moderate systematic difference in the extent of quality variation in the several series with Series A most restricted in quality variation, Series B somewhat more extended, and Series C somewhat further extended. The desire was, however, to keep all three distributions within the bounds of usual expected variability for color-print inspection and amateur photography. Then the obtained correlations, which on statistical principle might be expected to increase with range of variation, all might be considered conservative with respect to the situation of interest.

The attempt to control the ranges was made in a preliminary experiment in which 5 *O*s sorted many prints into five categories: 'like very much,' 'like somewhat,' 'indifferent,' 'dislike somewhat,' and 'dislike very much.' Then, by reference to the observed frequencies of these sortings three series of prints were so selected as to represent successively greater quality spreads. As it turned out, the differences in variability were insufficient to produce consistent differences in the correlations; and in view of the small *n* this is not surprising.

Procedure. All preferential judgments were made individually at a gray-topped table under about 45 footcandles of 4500°K white fluorescent light. Every *O* served in two sessions, one session for each method.

In case of the method of single stimuli, a double slot-box was used for sorting into the categories 'like' and 'dislike.' *E* placed a series of prints in a special box from which *O* could conveniently withdraw them one at a time, and instructed him to "please sort these Kodacolor Prints one at a time into the appropriate slots, according to your like or dislike for the color balance and density. When in doubt, make a best guess."

In case of the method of paired comparisons, the slot-box was replaced by a small easel upon which *E* juxtaposed the members of each pair of prints for comparison. The instruction was to "please indicate which of these two prints you prefer on the basis of color balance and density. When in doubt, make a best guess."

All the *O*s made five replications in all series with the method of single stimuli. The object of this was to secure the same amount of data by this method as was secured by the method of paired comparisons. Since the space error was not compensated in the method of paired comparisons, only 55 observations (11 prints $\times$ 5 replications) per *O* per series were required for this purpose.

The orders of both the print series and the experimental sessions were randomized among the *O*s. Every *O* completed all three series by a given method in a given session.

Results. The frequencies of the *O*'s preferences for all prints in all series by both methods were tabulated as a preliminary to correlation. These tabulations were carried out for the first trial alone, and for the first two trials

⁵ Dr. R. W. Burnham and Mrs. Margaret Duck assisted substantially in the conduct of this study.

together, as well as for the summation of all five replications. Thus, it would be possible to see what effect, if any, is produced on the correlation by the amount of data taken by the method of single stimuli. In the case of the five replications (with $n = 11$), the amount of data is the same, in the case of the two replications it is 0.4 as much, and in the case of the single trial it is only 0.2 as much as by the method of paired comparisons.

The Pearson correlations uncorrected for attenuation are given in Table I and show that even with the single trial the method of single stimuli yielded much the same results as the method of paired comparisons. Furthermore, plots of the correlation diagrams indicated that the relationship between the two variables was essentially linear.

Guilford's work had already shown that subjecting the observed frequencies to elaborate scaling routines of the Thurstone type would be

TABLE I

CORRELATIONS BETWEEN THE METHOD OF SINGLE STIMULI AND THE METHOD OF PAIRED COMPARISONS

Series	Replications by method of single stimuli		
	1	2	5
A	.944	.962	.980
B	.963	.969	.960
C	.969	.970	.968

superfluous and would not affect the results.⁶ This is because the observed frequencies are themselves in relative scale form.

The above results showed such consistently good agreement between the two methods, that the suggestion arose that the performance of the *Os* with the one method might be influencing systematically their performance with the other method. Such spurious interdependence of results seemed unlikely for several reasons: (1) Since the observations by the two methods were a day or more apart, there would be difficulty in recalling the performance by the first method used. (2) The fact that all three series by a given method were run in the same session, would make doubly difficult keeping unscrambled any effective memories of performance in particular series. (3) There seemed little motivation to remember, especially since the *Os* were not told that methods were to be compared.

Nevertheless, the problem seemed important enough to eliminate any

⁶ J. P. Guilford, The method of paired comparisons as a psychometric method, *Psychol. Rev.*, 35, 1928, 494-506; Some empirical tests of the method of paired comparisons, *J. Gen. Psychol.*, 5, 1931, 64-77.

possibility of interdependence because of the same *O*s using both methods. Therefore, an entirely different group of 12 *O*s was asked to repeat the observations with the method of single stimuli and two replications. Incidentally, *E* also was different. When these results were correlated with those from the method of paired comparisons by the original 12 *O*s, the coefficients shown in Table II were obtained. Despite the expected losses due to the differences in *O*s, the Pearson correlations generally remained above 0.80. In Series C, where the highest correlations might be expected, the coefficients were above 0.97. Still higher correlations would result if the correction for attenuation were made. These results confirm the view

TABLE II
CORRELATIONS BETWEEN THE METHOD OF SINGLE STIMULI AND THE METHOD OF
PAIRED COMPARISONS WITH DIFFERENT *O*s FOR EACH METHOD

Series	Replications by method of single stimuli		
	1st	2nd	Both
A	.831		
B	.798	.771	.806
C	.990	.875	.844
		.975	.986

that the high correlations in the original comparisons in Table I are not an artifact of memory.

PART II

This part of the study was quite similar to the first except that the judgments were of projected color transparencies rather than prints, and the *O*s worked in groups rather than individually. There was some reason to anticipate lower correlations with the transparencies because of the general reduction of reference colors in the darkened projection-room.

Choice of materials. Three series of 11 Kodachrome transparencies each were assembled from various sources. A similar though more careful attempt was made than in Part I to secure a moderate systematic difference in the quality ranges of the several series. In this case 15 *O*s participated in a preliminary evaluation of 179 slides. The same five quality categories as before were employed in this evaluation, and the three series to be used experimentally were selected on the basis of the observed frequencies of these sortings.

Procedure. All judgments were made in a dim projection-room by means of the Master Model Kodaslide projector producing a screen brightness of 165 ft.-L with empty slideholder. The slide projections were about 2.5×4 ft. in size and viewed at a distance of 16 to 21 ft. Sharp focus was maintained continually by an assistant a few feet from the screen using servo remote control. In case of the method of paired comparisons, a pair of matched Master Model projectors permitted simultane-

ous projection. The instructions given the *O*s were like those in Part I; but every *O* recorded his own judgments by making check marks under appropriate headings on a record sheet attached to a clipboard.

In Part I it had been found that five replications by the method of single stimuli gave almost the same result as two replications; therefore only two replications were used in Part II. Every *O* served in four sessions, one session for the method of

TABLE III

CORRELATIONS BETWEEN THE METHOD OF SINGLE STIMULI AND THE METHOD OF PAIRED COMPARISONS

Series	Replications by method of single stimuli		
	1st	2nd	Both
A	.940	.904	.950
B	.927	.935	.934
C	.948	.936	.948

single stimuli and three sessions for the method of paired comparisons. Half of the *O*s worked with one method first and the other half with the other method first.

Results. As in Part I, the Pearson correlations were computed between the observed preferences for all prints in all series by both methods of judgment. Since there were available two replications or trials by the method of single stimuli, the computations could be carried out for the first trial alone, the second trial alone, and both trials together. These correlations, unadjusted for attenuation, are presented in Table III; and repre-

TABLE IV

CORRELATIONS BETWEEN THE METHOD OF SINGLE STIMULI AND THE METHOD OF PAIRED COMPARISONS WITH DIFFERENT *O*s FOR EACH METHOD

Series	Replications by method of single stimuli		
	1st	2nd	Both
A	.869	.909	.914
B	.777	.788	.786
C	.929	.896	.924

sent the results secured with one and the same group of *O*s using both methods of judging.

When judgments by the method of single stimuli were secured by an entirely different group of *O*s and these frequencies were correlated with the paired comparison frequencies of the original group, the results were as shown in Table IV.

These correlations are of much the same order as those in Part I. The correlations do run somewhat lower in Part II; but there would seem no more chance of a difference of this degree being due to the change from

color prints to projected transparencies than to the shift from individual to group judging.

Evidently there is substantial agreement among the present results with the color prints and transparencies and the experimental results cited from the literature. All of these studies have been concerned with preferential judgments of the experimental-esthetic type. These facts, in conjunction with the apparent similarities in the methods themselves, suggest that the high correlation between the methods may be rather general for judgments of this type.

Of course, this does not mean that the method of single stimuli can be recommended unreservedly for all phases of subjective, qualitative determinations. Indeed, there may be numerous situations where neither method is applicable. Thus, if there are more than a few items to be compared, the method of single stimuli would seem to be indicated; but, on the other hand, if the differences between items are too small to be perceived in the absence of juxtaposition, the method of paired comparisons may seem necessary. Close comparisons of sharpness or graininess, for instance, might require the latter method.

SUMMARY

(1) A group of 12 *O*s made preferential judgments of three separate series of color prints, both by the method of single stimuli and the method of paired comparisons. In all three series the original uncorrected product moment correlations of the results by the two methods exceeded 0.95. Even when the two different methods were used by two different groups of 12 *O*s, the correlations ranged from 0.77 to 0.99. Therefore, in situations, like the judging of color prints, where either method is applicable, much the same results may be expected from the method of single stimuli as from the method of paired comparisons.

(2) A different group of 12 *O*s made preferential judgments of three separate series of projected color transparencies, using the same two methods. In all three series the correlations were 0.90 or more; and when the two methods were used by different groups of *O*s the correlations ranged from 0.78 to 0.93. Therefore, in situations like the judging of projected transparencies, where either method of judging is applicable, much the same results may be expected.

(3) The correlation between the methods proved to be as high when only one-fifth as much data was taken by the method of single stimuli, viz., a single trial, as when the same amount of data was taken by both

methods. The relatively easy method of single stimuli appears to be the more efficient method.

(4) Independent experiments of three other investigators with other materials showed the same close agreement between the methods. The data as a whole, in conjunction with apparent similarities in the methods themselves, suggest that the agreement may be fairly general for preferential judgments of this type.

(5) There are situations, of course, in which only one, if either, method is applicable. Thus when there are only a few items which are so similar that they must be juxtaposed to detect differences, the method of paired comparisons still seems indicated. Whether there are few or many items which do not require juxtaposition, the method of single stimuli seems preferable.

THE LOWER THRESHOLD OF VISIBLE MOVEMENT AS A FUNCTION OF EXPOSURE-TIME

By ROBERT H. BROWN and JACK E. CONKLIN,
Naval Research Laboratory, Washington, D.C.

The problem of this experiment is to determine how the lower threshold of movement varies with exposure-time. At very slow rates movement can only be inferred upon the basis of observed changes in position after intervals of time; as for example the hour hand of a clock. As the speed of movement increases, threshold is reached above which a continuous change of position in space can be observed. With further increases in speed, the discrimination of movement again fails at an upper threshold.

Careful control and analysis of the effects of interacting variables are essential in determining the lower thresholds.¹ In addition to exposure-time, the important variables are the distance the stimulus-object moves, its retinal position, the intensity of illumination, distance of observation, form and size of the stimulus-object, structure and dimensions of the field, conditions of fixation, and state of light or dark adaptation. Of these factors, the field is an especially important variable as it may interact with the exposure-time.

In measuring the lower threshold, various investigators have attempted, by restricting the field, to eliminate inferences of movement. Aubert, in the first experiments on the lower threshold, determined the effect of surrounding objects.² When *S* fixated a moving line on a rotating drum and had a full view of stationary parts of the apparatus, the threshold was of the order of 1' to 2' of visual angle per second. When the apparatus was so concealed that *S* could see only the drum, a tenfold increase in the threshold resulted. Similar increases were obtained when the field was limited to a small point in a circle.³ When, however, Aubert eliminated the station-

*Accepted for publication January 28, 1953. The opinions or assertions contained in this paper are those of the authors. They are not to be construed as official or reflecting the views of the Navy Department or the Naval Service at large.

¹J. F. Brown, The thresholds for visual movement, *Psychol. Forsch.*, 1931, 14, 249-268; J. L. Kennedy, The nature and physiological basis of visual movement discrimination in animals, *Psychol. Rev.*, 43, 1936, 494-521; C. H. Graham, Visual perception, in S. S. Stevens (ed.) *Handbook of Experimental Psychology*, 1951, 868-920.

²Hermann Aubert, Die Bewegungsempfindung, *Arch. f. d. ges. Physiol.*, 39, 1886, 347-370.

³Klara Grim, Über die Genauigkeit der Wahrnehmung und Ausführung von Augenbewegungen, *Zsch. f. Sinnesphysiol.*, 45, 1910, 9-26.

ary basis entirely by presenting a point of light in a dark room, the threshold became highly variable. Frequent reports of illusory movement—Charpentier's illusion—were reported.

The only systematic determinations of the lower threshold of movement for different stimulus-durations have been made by Dimmick and Karl.⁴ In their experiment, a moving white dot was presented in front of a uniform, black screen. S reported the direction of the dot's movement when the screen and dot were illuminated for varying intervals of time. Under these conditions, the threshold decreased very rapidly at first and then more slowly as the exposure-time was increased from 0.5 to 4 sec. For exposure-times of 0.5 and 1 sec., the distance traversed by the spot at a threshold-rate is constant and of the order of 1' of arc at S's eye. For longer exposures, an increasingly greater distance is required for the discrimination of movement.

The present study represents a further attempt to determine the influence of exposure-time upon the lower threshold. Conditions were obtained that exclude inferences of movement based upon visual comparisons of position. The essential features of our method are (1) provision for initial fixation of a moving spot, (2) presentation of the small-spot of light in the dark, and (3) determination of the lowest rate at which the direction of the movement can be discriminated. The lower threshold was measured under these conditions for six stimulus-durations.

Method. Observations were made in a dark room. Before every trial S regarded the fixation-point on an illuminated screen which was directly in front of him. During a trial he observed a spot of light projected on the screen. His head was immobilized by means of a biting board. He communicated telephonically with E who was in an adjoining room with the equipment for controlling the rate, duration, and direction of spot's movement. When S perceived movement, he reported its direction by pressing one of two switches, held in his hands, which turned on small lights in E's room. His reports were recorded as 'right' or 'wrong' and his lower threshold was computed by the method of constant stimuli.

Apparatus. As the apparatus used in this study has been separately described,⁵ a brief description here will suffice. A white cardboard screen, (30 × 36 in.), on which the moving spot was projected, had a green fixation-point ($\frac{3}{4}$ in. in diameter) at its center. The screen was illuminated from the top by shield lamps. Its brightness at the fixation-point, as measured by a Macbeth Illuminometer, was 0.021 ml.

The moving spot, a green circular area of the exact size of the fixation-point, was projected onto the screen from an oscilloscope. Its brightness was 0.026 ml.; approximately the same as that of the screen. The oscilloscope, placed above and behind S, was so adjusted that the spot of light, when turned on, would fall upon the fixation-point. A relay, activated by a timing device, extinguished the lamps

⁴ F. L. Dimmick and J. C. Karl, The effect of the exposure time upon the R. L. of visible motion, *J. Exper. Psychol.*, 13, 1930, 365-369.

⁵ J. E. Conklin, A. Baldwin, and R. H. Brown, Apparatus for measuring the threshold for the visual discrimination of direction of movement, this JOURNAL, 66, 1953, 289-294.

illuminating the field at the same time that the moving spot was projected upon the screen; thus the screen and the fixation-point disappeared with the appearance of the stimulus-spot.

The direction and rate of movement of the stimulus-spot were controlled by the polarity of the oscilloscope and the changes of voltage across its horizontal deflection plates, respectively. The rates of movement in minutes of visual angle per second varied by steps of 3.22' between 1.61' and 27.37'. They were 1.61', 4.83', 8.05', 11.27', 14.49', 17.71', 20.93', 24.15', and 27.37'.

The exposure-times, regulated by a Bodine synchronous motor and various cams, were 0.51, 0.98, 1.97, 3.99, 8.01, and 16.00 sec.

S's observations were made monocularly at a distance of 72 in. from the screen.

Subjects. The Ss, enlisted men from the U. S. Navy assigned to serve in the study, were four in number. Though naïve psychologically and inexperienced in making psychophysical judgments, they were highly cooperative and interested in their work throughout the course of the investigation.

Procedure. S sat in the dark for 3 min. while E read the instructions to him over the intercommunicating telephone.⁶ The screen was then illuminated for 6 sec., during which S focused upon the fixation-point. Then the stimulus-object, the moving spot, appeared at the fixation-point as the lights illuminating the screen were simultaneously turned off. After the exposure, during which the stimulus-object was moved, S reported the direction of the movement. The screen was then illuminated for 19 sec. after which the stimulus-object was again presented at the same exposure-time but at a different rate of movement. This procedure was repeated for the nine different rates of movement at every one of the six exposure-times.

During the first half of an experimental session, there were six series of presentations in which the stimulus-object moved at every one of the nine rates. S then rested for as long a period as he desired and the trials were resumed. The second half was a repetition of the first except that the random order determined for the first half was reversed.

The rates and durations of movement in the trials were determined before a session on the basis of a table of random numbers. The direction of the movement, which was horizontal, was selected by chance, subject to the following restrictions: (1) an equal number of left and right movements should occur in each half of the experimental session; and (2) no more than three consecutive movements should be made in the same direction.

Since the Ss reported vertical components in the movements, which were illusory (autokinetic) as none was objective, it was deemed desirable to determine whether there was an illusory component in the horizontal plane. Accordingly, at the close of the main experiments, the Ss were informed that they would be given a series of trials in which the spot moved at very very slow rates. They were instructed to do the best they could in determining the direction of its movement. A stationary spot was then presented 48 times at each of the stimulus-durations used in the main experiments.

Comments by the Ss regarding the objective or illusory nature of the movements they observed and the bases they used in judging direction were noted throughout

⁶ For the instructions, see Conklin, Baldwin, and Brown, *op. cit.*, 293.

the experiment. In addition, at the end of the experiment, the Ss were asked to indicate their confidence, or lack of it, in their judgments, to describe movements in any direction other than the horizontal, and to comment upon the reality of the non-horizontal movements.

Results. The experiment yielded 48 judgments from every S regarding the movement of the stimulus-spot at every one of the nine rates of movement and six exposure-times employed. These judgments were grouped accordingly as they were 'right' or 'wrong' and the percentages 'right' were computed for every rate of movement and every exposure-time.

The proportion of correct judgments increases with rate of movement from 1.61' to 27.37' of visual angle per second. These data were corrected

TABLE I
RL OF MOVEMENT IN MINUTES OF VISUAL ANGLE PER SECOND
Exposure-Time (in sec.)

S	0.51	0.98	1.97	3.99	8.01	16.00
A	14.22	12.32	11.91	11.42	12.12	12.53
B	22.43	16.15	12.92	9.23	6.35	6.78
C	16.22	13.01	12.38	12.38	9.57	9.72
D	17.87	14.43	11.50	10.35	8.62	6.70
Av.	17.68	13.98	12.18	10.85	9.17	8.93

for guessing and examined for normalcy. Since it may be safely assumed that S's guesses would be right as frequently as wrong, we corrected the results for guessing by subtracting the proportion 'wrong' from the proportion 'right.' When these proportions were plotted against rate of movement for the various exposure-times, a family of ogive curves resulted. A plot of the average of 10 distributions with comparable medians and ranges, made on probability and log probability paper, indicates an approximately normal distribution with reference to rate but not with reference to the logarithm of the rate.⁷ The lower threshold of visible movement may therefore be regarded as the arithmetic mean of the corrected frequency distribution.

From the proportion of correct judgments for every exposure-time, thresholds were computed by Spearman's mean method.⁸ Table I gives the thresholds obtained for the individual Ss and the average of them all. It

⁷ This finding, which indicates that the phi-gamma hypothesis applies to the lower rate threshold rather than the phi-log-gamma hypothesis, agrees with an earlier observation by Dimmick and Karl, *op. cit.*, 367.

⁸ Arithmetic means of the distributions were computed with the formula given by R. S. Woodworth, *Experimental Psychology*, 1938, 400-403.

shows considerable variability among the *Ss* as well as scatter for each *S*. This variability necessitated a statistical analysis of the results. Since the experimental procedure was appropriately randomized and replicated, an analysis of variance was made to determine the reliability of differences in the thresholds attributable to exposure-time and to *S*.

TABLE II
ANALYSIS OF VARIANCE FOR THE RATE-THRESHOLDS

Source of variation	df	Mean square	F-ratio
Exposure-times	5	43.95	7.35*
<i>Ss</i>	3	0.86	0.14
Error	15	5.98	

* Significant at 1-% level.

Table II presents the results of this analysis. It shows that the *Ss* constitute a homogeneous group and contribute little to the variation in the thresholds, and that the exposure-times provide a very significant source

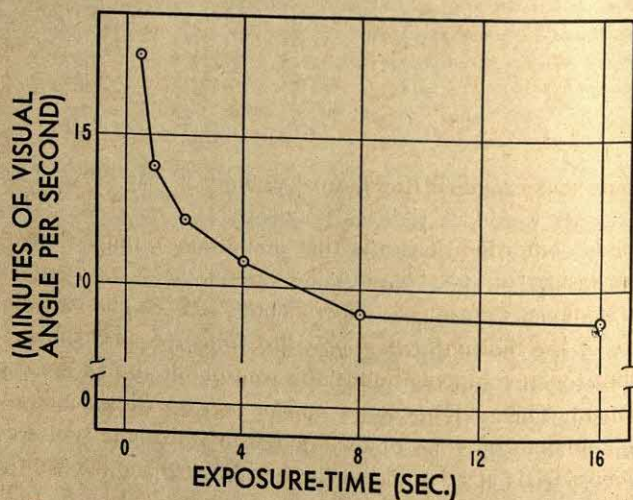


FIG. 1. THE RL OF VISUAL MOVEMENT AS A FUNCTION OF EXPOSURE-TIME

of variation. Individual *t*-tests, based on only three degrees of freedom, show a significant difference at the 5-% level of confidence between the thresholds for 0.51 and 0.99 sec., but not between any two of the adjacent thresholds for the other exposure-times.

Fig. 1 presents the average of four *Ss*' thresholds for rate of movement. The curve indicates in general that the lower threshold for visual move-

ment decreases with an increase in exposure-time. The rate of decrease is rapid for shorter exposures, and the curve appears to approach an asymptote of a rate equal to about 9' of visual angle per sec. for exposure-times greater than 16 sec.

The Ss' reports of movement for a stationary spot are given in Table III. The observed number of movements to the right, which are entered in the table, may be compared with the 24 responses expected in terms of random horizontal motion. A chi-square test for all Ss combined shows no systematic bias from chance. Individual chi-square tests indicate one S (B) had a tendency to report more movements to the right than to the left. We may conclude from these data that autokinetic movement was

TABLE III
NUMBER OF JUDGMENTS OF MOVEMENT TO RIGHT FOR THE STATIONARY SPOT
Exposure-Time in Seconds

S	0.51	0.98	1.97	3.99	8.01	16.00	Chi Square
A	17	23	21	20	21	24	7.02
B	27	36	24	32	23	22	18.52*
C	26	29	22	27	24	24	3.52
D	28	27	27	19	23	21	5.78
Total	98	115	94	98	91	91	7.76

* Significant with five degrees of freedom at 1% level.

random under our conditions and that it did not introduce a systematic error in our results.

In their comments during the experiments, the Ss emphasized the objective nature of the movement they observed. This movement included not only the horizontal component under E's control but also an illusory vertical component. Their reports at the end of the experiment indicated that they were confident of 90% of their judgments, that the spot seemed to move on most trials at an angle to the horizontal plane but with a trend right or left, that the vertical movement was controlled by E, and that it was not illusory. In general, the comments of the Ss suggested that they based their judgments upon an observed continuous change in position.

Discussion. The results may be summarized as follows. (1) The lowest rate of movement required for the judgment of visual movement decreases as the exposure-time increases. (2) Autokinetic movement was random in direction under the conditions of this experiment.

The first observation agrees with the relationship found by Dimmick

and Karl for a white dot moving in front of an illuminated screen.⁹ Restriction of the field of view in the present experiment to a moving spot in the dark serves to raise the threshold. Moreover, the extent to which elimination of a background raises the threshold is dependent upon the exposure-time. At a 0.5-sec. exposure the threshold for the movement of a spot in the dark is approximately seven times greater than the threshold measured by Dimmick and Karl with a visible background. The threshold is approximately 15 times greater at a 4-sec. exposure which permits the S to utilize visual cues indicating a change in the position.

That the autokinetic movements were not constant in any one direction may be due to our experimental conditions: to the fact that the stimulus-object was moved equally often to the right and to the left, and that similar instructions and reinforcements for judgments of movement in both directions were given the Ss. Autokinetic movements prevailing in one direction would doubtlessly increase the threshold for movement; they certainly would, as Aubert pointed out,¹⁰ render judgments of movement difficult when a spot of light is used in darkness.

Conclusions. The results of this study indicate (1) that the threshold of visible movement decreases (becomes better, more acute) as the exposure-time is increased; (2) that the decrease, large at first as the short exposure-times are lengthened by equal steps, becomes progressively less as the exposure-times become longer; and (3) that autokinetic phenomena did not under the conditions of this study affect either the observations or the reports of our Ss.

⁹ Dimmick and Karl, *op. cit.*, 368.

¹⁰ Aubert, *op. cit.*, 347-370.

CERTAIN TEMPORAL CHARACTERISTICS OF THE RECALL OF VERBAL ASSOCIATES

By W. A. BOUSFIELD, C. H. W. SEDGEWICK, and B. H. COHEN,
University of Connecticut

This paper deals with an interpretation of certain temporal characteristics of recall following varying amounts of reinforcement of the learning of a list of words. The method employed is based on the supposition that if a rational equation can describe adequately the relationship between an independent and a dependent variable, we can assume support for the assumptions incorporated in the equation. This type of procedure is circular in the sense that the data furnish cues for the choice of assumptions, which are in turn tested by seeing how well their mathematical expression fits the data.

From a theoretical point of view, our technique of analysis is closely related to an earlier study by Bousfield and Sedgewick who used the simple procedure of asking subjects to list items in a specified category, *e.g.* animals, names, or U.S. cities.¹ Plots of the cumulative totals of these items as a function of time spent in their production were described adequately by the following modified exponential equation:

$$n = c (1 - e^{-mt})$$

In this equation, n = the total number of items produced within time t ; c = a constant, the asymptote approached by the function, and defined rationally as the total supply of items available for recall; m = a constant, the slope factor. The rate of depletion of the supply of items is given by the differential equation:

$$dn/dt = m (c - n)$$

We may see from this equation that the rate at which items are produced is equal to the constant m times the number of the remaining items available for recall, $(c-n)$. Johnson, Johnson, and Mark have confirmed and extended this analysis.² We are indebted to them for calling attention to the fact that the rate refers to depletion rather than to the emission of the supply of items.

* Accepted for publication January 4, 1953. This is Technical Report No. 3 under contract Nonr-631 (00) between the Office of Naval Research and the University of Connecticut.

¹ W. A. Bousfield and C. H. W. Sedgewick, An analysis of sequences of restricted associative responses, *J. Gen. Psychol.*, 30, 1944, 149-165.

² D. M. Johnson, R. C. Johnson, and A. L. Mark, A mathematical analysis of verbal fluency, *J. Gen. Psychol.*, 44, 1951, 121-128.

The specific purpose of this study is to apply a similar analysis to a situation in which groups of Ss, given varying numbers of presentations of a list of stimulus-words, were then asked to recall as many as they could. We were specially concerned with the testing of the conclusion, reached in the earlier study, that the rate of production of items at any time is a function of the number of remaining items available for recall.

The data used in the present analysis were obtained from an earlier study designed for the purpose of investigating the occurrence of sequences of related words during recall.³ We have employed these data because they yielded a family of empirical curves of recall whose experimentally variable parameter was the number of repetitions of the list of stimulus-word during the learning period. It was our supposition that if a rational equation could describe such a family of recall curves, we would expect that the variations in the constants of the equation should relate meaningfully to the parameter of the degree of reinforcement occurring in the learning period.

METHOD

(1) *Subjects.* The Ss were five groups of undergraduate students enrolled in the laboratory sections of an introductory course in psychology. The sizes of these groups were as follows: Group I had 49 Ss; Group II, 52; Group III, 46; Group IV, 49; and Group V, 47.

(2) *Stimulus-materials and apparatus.* The list of stimulus-words comprised 60 nouns, 15 in each of the following four categories; namely, *animals*, *names*, *professions*, and *vegetables*. The words in each category were matched with respect to syllables and their Thorndike-Lorge frequency-of-usage counts. Five separate randomizations of this list were prepared, the first of which was as follows:

muskrat, blacksmith, panther, baker, wildcat, Howard, Jason, printer, chemist, radish, mushroom, Otto, plumber, pumpkin, chipmunk, Amos, Wallace, parsnip, milkman, druggist, leopard, woodchuck, Adam, grocer, Simon, Owen, lettuce, giraffe, turnip, garlic, rhubarb, typist, eggplant, Noah, zebra, donkey, Gerald, dentist, otter, parsley, spinach, Oswald, weasel, broker, waiter, florist, Bernard, dancer, reindeer, Byron, cabbage, melon, badger, mustard, diver, carrot, Sherman, camel, baboon, Moses.

Slides were prepared for each of the five randomizations of these words. To expose them on a screen, we used a projector with a mask which could be moved from word to word by *E*, and in time with the flash of a small light occurring at 3-sec. intervals. The light was mounted on the projector but was not visible to the Ss. In this way the words were exposed singly and for intervals of 3 sec. Mimeographed sheets, $8\frac{1}{2} \times 11$ in. in size, were prepared for use by the Ss in writing

³ Bousfield and B. H. Cohen, The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates, Technical Report No. 1 under Contract Nonr-631 (00) between the Office of Naval Research and the University of Connecticut.

the words they were able to recall. These sheets were divided by vertical lines into five columns which were numbered in order from one to five.

(3) *Procedure.* The general procedure was to give the *Ss* varying numbers of presentations of the stimulus-words during a single learning period. Group I received a single presentation of the words; Group II, two presentations; Group III, three presentations; Group IV, four presentations; and Group V, five presentations. Duplication of a serial order was avoided in all cases of multiple presentations of the words by having prepared the five separate randomizations of the words.

In conducting the experiment, *E* distributed the mimeographed sheets to the *Ss* and informed them that a series of words would be projected on the screen, after the exposure of which, they were to write as many of the words as they could recall.

TABLE I

MEANS AND SD OF TOTALS OF WORDS RECALLED AS A FUNCTION OF TIME SPENT IN RECALL FOR FIVE LEVELS OF REINFORCEMENT

Time (in min.) after starting signal

Rein- force- ments	<i>Ss</i>	2		4		6		8		10	
		M	SD	M	SD	M	SD	M	SD	M	SD
1	49	16.18	3.65	19.43	4.33	21.88	5.35	22.96	6.06	23.96	6.69
2	52	18.19	4.32	23.48	5.06	26.12	5.49	27.83	5.97	29.12	6.30
3	46	20.72	4.81	26.54	6.04	28.85	6.21	30.37	6.13	31.70	6.23
4	49	21.41	4.42	27.84	6.51	31.04	6.85	33.18	7.00	35.12	7.00
5	47	22.21	5.16	30.55	6.76	34.26	6.50	36.26	6.90	37.87	7.31

He explained further that in recall they were to start by writing the words in the first column of the data sheets. At periodic intervals they would be asked to shift to the next consecutive column and continue with more words. *E* then exposed the words. For all multiple projections of the words, a rest-period of 30 sec. was allowed between successive presentations of the list. The signal to start writing words was given 3 sec. after the exposure of the final word. The signals to shift to succeeding columns and continue with more words were given at 2-min. intervals throughout the 10-min. period allowed for recall.

RESULTS AND DISCUSSION

The first step in the preparation of the data was to compute the cumulative totals of words recalled for each of the consecutive 2-min. periods. This was done for every *S*. We then computed the means and standard deviations of the cumulative totals of words recalled for every one of the five groups of *Ss*, *i.e.* for each of the five levels of reinforcement. These data appear in Table I. The means from Table I are shown by the dots in Fig. 1. The curves were derived from our rational equation which we shall presently explain. Inspection of these data reveals the expected increase in the number of items recalled as a function of the degree of reinforcement. This relationship persists throughout the period of recall.

We now consider the derivation of our rational equation. To begin with, the five plots of the means of the cumulative totals of our data were sufficiently similar to justify the assumption of a family of curves describable in terms of a single equation whose constants should vary systematically with the degree of reinforcement. To generate such an equation it was necessary to start with a group of reasonable assumptions. It was reasonable to assume that the supply of associates for the Ss was limited. Every curve

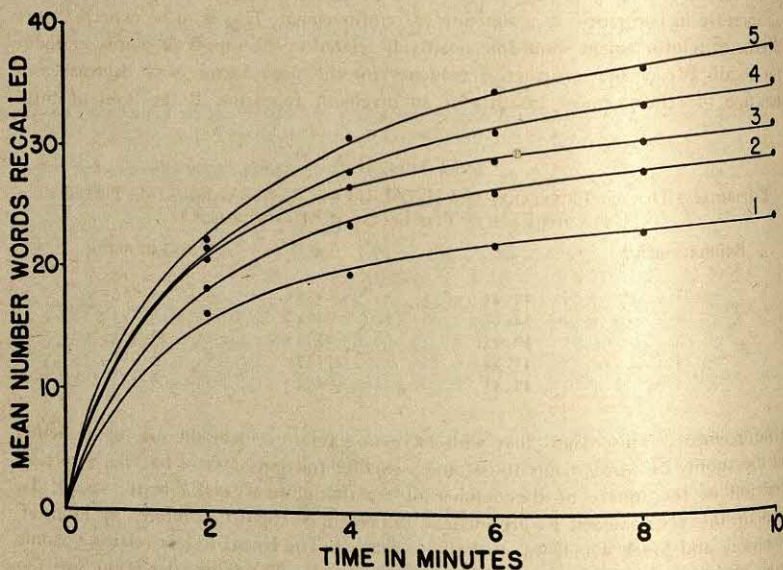


FIG. 1. CURVES OF RECALL FOR VARYING DEGREES OF REINFORCEMENT
The curves are numbered according to the number of presentations of the stimulus-list during learning. The dots represent the empirical means given in Table I.

would therefore approach a horizontal asymptote, which in turn would be represented by a constant in the equation. On the basis of the work of Bousfield and Sedgewick it was also reasonable to assume that the rate of approach to the asymptote at any time should be a function of the total number of items available for recall.⁴ We therefore undertook to fit our data to the modified exponential equation proposed by Bousfield and Sedgewick. It became obvious, however, that our data could not adequately be described by this equation. The next step was to explore the possibility of the derivation of other equations which would preserve our basic assumptions and at the same time fit our data.

⁴ Bousfield and Sedgewick, *op. cit.*, 149-165.

In this undertaking, we started with the Bousfield-Sedgewick differential equation described above. The characteristics of our data suggested that dn/dt might be proportional to the square of $(c - n)$, i.e. to $(c - n)^2$. The solution of this differential equation, satisfying the conditions that $n = 0$ when time (t) also equals 0, gives the hyperbola:

$$n = c^2 mt / 1 + cmt$$

This equation provides a family of theoretical curves which fit our data well as shown by the curves in Fig. 1. We obtained the constants c and m for these curves by a method of averages. They are given in Table II. Inspection of these data reveals an increase in constant c as a function of reinforcement. This is to be expected since degree of reinforcement should be positively related to the supply of words available for recall. We observe, further, a tendency for the slope factor m to decrease as a function of reinforcement in spite of an inversion appearing at the level of three

TABLE II

EMPIRICALLY DERIVED CONSTANTS AND HALF-LIVES FOR RATIONAL EQUATION FITTED TO DATA FOR EACH OF FIVE LEVELS OF REINFORCEMENT

Reinforcements	c	m	$t_{1/2}$ (in min.)
1	27.54	.0235	1.55
2	34.12	.0164	1.79
3	36.05	.0191	1.45
4	41.34	.0127	1.91
5	45.45	.0108	2.03

reinforcements. This means that with increasing reinforcement the rate of depletion of the supply of words tends to become a smaller fraction of $(c - n)^2$, i.e. a smaller fraction of the square of the number of remaining items available for recall. To explain this phenomenon we are inclined to favor the supposition offered by Johnson, Johnson, and Mark which was that the constant m may be negatively related to some type of interference effect taking place during recall.⁵ The larger the m the less the interference. We do not like the connotations of the term 'interference,' but prefer to use it rather than complicate our discussion by the introduction of a new term. Other factors being equal, the larger the supply of items the greater the interference. It may be noted that Johnson, Johnson, and Mark supported the assumption of interference on the basis of a significant negative correlation between the constants c and m . In the course of our analysis the question arose as to whether it might not be possible to derive a meaningful index based on a relationship between the constants c and m , since the rate of depletion of the supply of items is a function of both these constants. The concept of the half-life appeared to meet these specifications. In our terms the half-life would be the time when half of the supply of available items had been produced. As derived from our rational equation, the half-life ($t_{1/2}$) is given by the following equation:

$$t_{1/2} = 1/cm.$$

The values of $t_{1/2}$, in minutes, for each of our five levels of reinforcement are shown in Table II. The trend is toward a progressive increase, though we again

⁵ Johnson, Johnson, and Mark, *op. cit.*, 124 f.

encounter an inversion at the level of three reinforcements. The question now arises, whether the concept of the half-life may be regarded as a psychologically meaningful measure? We would propose tentatively that the half-life may be regarded as a special index of efficiency of recall. In our terms, the shorter the half-life the greater the efficiency of recall. The formula for the half-life shows that if the supply of items (c) were held constant, the greater the m , i.e. the less the interference effect, the greater the efficiency. Conversely if the interference effect were held constant, the greater the supply of items (c) the greater the efficiency. How are we to explain the apparent contradiction that with increasing reinforcement the S s produced items more rapidly but less efficiently? The answer is that we have not defined efficiency of recall here in terms of the simple rate of production of items. To estimate efficiency of recall, as we have defined it, we should know what the supply of items available for production is and we should also know how much interference is taking place during their production.

We shall consider the question of possible reasons to account for the difference between the equation we have proposed for the data of this study and the one used in the earlier studies of Bousfield and Sedgewick, and Johnson, Johnson, and Mark.⁴ More specifically, we should account for the difference between the two differential equations proposed.

As already indicated, both of these equations state that the rate of production of items is a function of the remaining available supply of items ($c - n$). The fact that in the case of the present experiment the supply ($c - n$) is squared may be taken to mean that the potency of this supply to determine rate is greater than in the previous experiment where the supply ($c - n$) was not squared. Furthermore this potency, when ($c - n$) is squared, is much greater when time (t) is small than when it is large. Parenthetically it may be noted that $(c - n)^2$ in our differential equation should be regarded for the time being as an approximation. On theoretical grounds there is reason to expect that the power to which ($c - n$) should be raised may well turn out to depend both on the initial habit strengths of the words in the stimulus-list and on the degree of their experimental reinforcement.

Let us consider the difference between the production curves for the two types of experimental situations. In the earlier studies, where the S s were asked to list items in a single specified general category, and without reinforcement of the items during the experimental situation, the production curves showed a relatively slow initial rise. In the present study where we obtained recall following experimentally induced reinforcement of a list of items, the rise in the production curves was initially relatively rapid as shown by the curves in Fig. 1. There is some justification for regarding the reciprocals of the latent periods of the items as indices of their habit strengths. Suppose now that we apply this assumption to our production curves of items. The latent periods of the items would be the times for their occurrence measured from the start of the period of recall. For our purposes, we may obtain the latent period, measured on the horizontal axis from the origin of a curve, for each decile of items. In this way we would obtain a table of progressively increasing latent periods for equal fractions of the total number of words recalled. The reciprocals of these latent periods would show a progression of decreasing habit strengths for the production curve we are considering.

⁴Bousfield and Sedgewick, *op. cit.*, 149; Johnson, Johnson, and Mark, *op. cit.*, 121.

Let us see what this analysis would show for two production curves, *A* and *B*. To make our illustration as simple as possible, we shall suppose that both curves approach the same horizontal asymptote. Curve *A* would represent data from the present study. It is marked by the fact that the rate of production of items which it portrays is initially relatively rapid. Curve *B* represents data obtained from our earlier studies, and is marked by the fact that the rate of production of items which it portrays is initially relatively slow. Using the method we have described we may obtain the progression of habit strengths for each decile of items for both curves. The result would be that the progression of magnitudes of hypothetical habit strengths for Curve *A* would show a slower initial drop than would be the case for Curve *B*. This in turn means that the initial habit strengths for Curve *A* are greater than the corresponding strengths for Curve *B*. We are thus able to account for the initially more rapid production of items in Curve *A* than in Curve *B*. It would also follow in this situation that the total amount of habit strength for Curve *A* would be greater than the total amount for Curve *B*. This relative potency on the part of the total strength of the supply of items represented in Curve *A* is what results in our finding it necessary to give $(c - n)$ a power greater than one in the *A*-type of situation, which is also the situation existing in the present experiment. We have still not answered the question of the essential difference between the types of situations which would account for the two types of production curves. It is our present supposition that the essential differences lies in the fact that whereas reinforcement was given to the words during the present experiment, no such reinforcement took place in the earlier studies where the *Ss* were simply asked to list items in a specified category. We may also note that in general the items having low habit strength would be expected to gain more from experimentally induced reinforcement than would be the case for items having relatively high habit strength. In so far as this effect operates, the result would be a less steep gradient in the resultant distribution of habit strengths, which is what we have regarded as the essential reason for the type of production curve obtained in the present study.

Our account has been highly speculative in character, and it is necessarily restricted to the results of our experimental conditions. We believe we have demonstrated, however, that certain temporal characteristics of the recall of verbal associates are a function of the reinforcement given these associates in an experimentally controlled situation. The use of rational equations to describe the production curves of verbal associates offers promising possibilities for the development of psychological theory.

SUMMARY

This study makes use of the data of an earlier experiment in which a list of 60 stimulus-words was presented in a learning situation once to one group of *Ss*, twice to a second group, three times to a third group, four times to a fourth group, and five times to a fifth group.⁷ Repetition of a serial order of the words was avoided in all cases of multiple presentations

⁷ Bousfield and Cohen, *op. cit.*

of the list. Immediately after learning, the Ss were allowed five successive two-minute intervals in which to write the words they were able to recall. The present study makes use of these data to determine the relationship between the cumulative totals of items produced and the time, up to ten minutes spent in recall, for each of the five degrees of reinforcement of the learning of the words. On the basis of what appeared to be tenable assumptions a differential equation was constructed to indicate the rate of depletion of the supply words, and from it was derived a hyperbolic equation to describe the functional relationship between the number of words produced and the time required for their production. These two equations are represented below as *a* and *b*:

$$\begin{aligned} * \quad \frac{dn}{dt} &= m(c - n)^2 \dots\dots\dots [a] \\ n &= c^2mt / (1 + cmt) \dots\dots\dots [b] \end{aligned}$$

The definitions of the terms of these equations are as follows: *n* = the total number of words produced in time *t*; *c* = the asymptote approached by the function, rationally defined as the total number of items available for recall; *m* = a constant (the slope factor), rationally defined as an inverse index of interference effect. The half-life, (*t*_½) for the production of words, interpreted here as an inverse index of efficiency of recall, is obtained by substituting appropriate values in the following equation: *t*_½ = 1/*cm*.

Attention is given in the paper to the rationale of the interpretations and their relationship to earlier research on the production curves of verbal associates.

LATENT LEARNING IN HUMAN SUBJECTS

By LEO POSTMAN and A. H. TUMA, University of California

In experiments on latent learning, which have focused on testing the reinforcement hypothesis, insufficient attention has been given to the question of what it is that the unrewarded *S* does or does not learn. Experimental tests of latent learning seem to have been based on the assumption that whatever the organism learns under normal conditions of reward it may also be expected to learn in the absence of reinforcement. The validity of this assumption is doubtful, especially in the case of irrelevant-incentive experiments.¹ It is quiet possible that in the absence of reward the *S* learns *some but not all* of the things he does in the presence of reward. In irrelevant-incentive experiments, for example, the *S* may learn that the irrelevant goal-object is present but may fail to acquire information about the path to the goal. In the critical test trials, such partial information would not be sufficient for a correct choice. Thus, conventional tests for latent learning may not be sufficiently sensitive, and it may be for this reason that they often yield negative results.²

For a full analysis of latent learning in irrelevant-incentive experiments, therefore, testing procedures at the end of training should be differentially sensitive to (1) *S*'s knowledge of the presence of the previously irrelevant goal-object and (2) *S*'s knowledge of the path leading to that object. In the present experiment, an experimental analogue of latent learning in human *Ss* was used to explore the effectiveness of such differential tests. The specific hypothesis under study was that *S* may learn about the presence of irrelevant incentives without learning their specific location.

METHOD

Learning task. The *S*'s task was to learn the correct pathway through a mental maze. Establishment of selective sets provided the analogue to the manipulation of drive. Stimuli conforming to *S*'s set served as relevant 'incentives,' stimuli falling outside the set as irrelevant 'incentives.'

The mental maze consisted of a series of 12 multiple-choice items. Each item consisted of four stimulus-words, typed side by side in solid capital letters. The words were numbered 1 through 4, with the numbers appearing above the words. The 12 multiple-choice items were arranged serially on a long strip of white cardboard. Additional strips of white cardboard were used to cover all parts of the maze except those to which the *S* was responding at any given moment.

* Accepted for publication December 12, 1952. The writers are indebted to Professor E. C. Tolman for his critical reading of the manuscript and for helpful suggestions.

¹ For an analysis of different types of latent learning experiment, see D. Thistlethwaite, A critical review of latent learning and related experiments, *Psychol. Bull.*, 48, 1951, 97-129.

² Thistlethwaite, *op. cit.*, 107-120.

The path through the maze was marked by words belonging to a specific meaning class, one member of this class appearing in each multiple-choice item. Two such paths were used in the course of the experiment, one marked by the names of fruit, the other by names of articles of clothing. The position of the correct word varied from item to item. The words, *fruit* and *dress*, respectively, marked the correct choices in the terminal positions.

Training procedure. There were two experimental groups—the *Transfer Group* and the *Recall Group*—and a *Control Group*.

The first task of all Ss was to learn a pathway through the maze to a criterion of two successive errorless trials. For half the Ss in each group, the correct path was marked by the names of fruit, for the other half by the names of articles of clothing. While learning this path, both experimental groups were exposed to the series of words marking the alternative pathway through the maze. For those learning the path marked by the names of fruit, the alternative path was marked by names of articles of clothing, and *vice versa*. For the Control Group, a series of unrelated words appeared in the positions corresponding to the alternative pathway.

The Ss were introduced to the learning task, and their selective set for a specific pathway was established, by informing them of the class of words marking the correct choices. The Ss were then presented with the numbers marking the first multiple-choice item, with the stimulus-words (and the rest of the maze) covered. As soon as S had made his choice, the words underneath the numbers were uncovered and S was requested to read the words aloud, at the rate of one word per second. If his choice was incorrect, S made another choice. The second choice was, of course, always correct since S recognized the correct word belonging to the critical meaning class. After S had chosen correctly, he moved on to the next item and the procedure was repeated for the rest of the items in the series.

It will be noted that S made his choices in terms of the numbers designating the four alternative words, *i.e.* he had to learn the *position* of the correct word at each choice point. Since S was required to read aloud *all* the words at each choice point, it is certain that experimental Ss had an opportunity to acquire information about the alternative pathway.

Tests for latent learning. Immediately upon reaching criterion, the Transfer Group and the Control Group were required to learn a second pathway through the maze. Again, a selective set was established by informing the Ss of the class of words marking the new pathway. This pathway was, of course, the one to which the experimental Ss, but not the control Ss, had been exposed during the initial training. The same procedure was followed as before, and learning was again carried to a criterion of two successive errorless trials.

The Recall Group was tested for retention of the words marking the alternative pathway. The meaning class which had marked the alternative pathway during training was indicated to the Ss and they were asked to write down as many of the words in this class as they could remember. These Ss were not asked to recall any of the other irrelevant stimulus-words, so that the transfer-test and the recall-test were conducted under the same selective set.

Two independent indices of latent learning were thus obtained. The performance of the Recall Group provided an estimate of the amount of incidental memory for irrelevant-incentive stimuli regardless of specific location. The comparison of the Transfer Group and the Control Group represents a conventional test for latent

learning, depending on the experimental Ss' incidental memory for the specific location of the irrelevant-incentive stimuli. The critical question is whether these two independent measures lead to different conclusions concerning the amount and significance of latent learning.

Subjects. The Ss were undergraduate students at the University of California. They did not know the purpose of the experiment and were assigned to the different conditions at random. There were 12 Ss in each of the three groups.

RESULTS

Original learning. Table I shows, for all groups, the average number of trials required to reach criterion and the average number of errors made in learning the first pathway. The three groups differed in speed of original learning, although analysis of variance of the two sets of scores failed to yield significant differences. For errors, the F -ratio was 2.90 ($df = 2$ and 32, $P > 0.05$); for trials, 3.21 ($df = 2$

TABLE I

AVERAGE NUMBER OF TRIALS TO CRITERION AND AVERAGE NUMBER OF ERRORS IN LEARNING ALTERNATIVE PATHWAYS THROUGH MENTAL MAZE

Group	First Pathway				Second Pathway			
	Trials		Errors		Trials		Errors	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Control	9.00	2.24	30.25	9.73	8.00	3.96	21.00	12.69
Transfer	10.17	3.40	29.42	8.86	9.08	3.71	21.83	8.48
Recall	12.25	3.34	38.08	9.22				

and 32, $P > 0.05$). Variations in speed of original learning were, nevertheless, taken into account in the analysis of the test-performances of the different groups.

Transfer test. A first measure of latent learning is provided by the difference between the Control Group and the Transfer Group in acquiring mastery of the second pathway through the maze. A significant difference in favor of the Transfer Group would constitute evidence for latent learning. As Table I shows, however, there is no evidence whatsoever for superior performance by the Transfer Group as measured by number of trials to criterion and by number of errors. The performance of the Control Group was actually a little better than that of the Transfer Group. It must be remembered, however, that the original performance of the Control Group was also somewhat better than that of the Transfer Group, and in evaluating the differences between the two groups, this variation must be taken into account. Both trial and error scores were, therefore, subjected to an analysis of covariance which showed the differences between the groups to be completely insignificant. For trials the F -ratio was 0.003; for errors, 0.036.

Fig. 1 shows Vincent learning curves of the Transfer Group and the Control Group for acquisition of the second pathway. The curves show the percentages of the total learning period (in trials) required to reach successive criteria.^a The form of the curve is quite similar for the two groups. A more sensitive index of

^a A. W. Melton, The end-spurt in memorization curves as an artifact of the averaging of individual curves, *Psychol. Monog.*, 47, 1936 (No. 212), 119-134.

latent learning may be provided by the number of errors on the *first trial* in learning the second pathway. This measure, however, again fails to reveal a significant difference in favor of the experimental group. For the Transfer Group, the average number of errors on the first trial was 7.50 with an *SD* of 1.89; for the Control Group 7.92, with an *SD* of 1.23.⁴ Analysis of covariance, taking account of variation in the number of errors in the original training, yields an *F*-ratio of 0.44. Clearly, there is no evidence for latent learning as measured by speed of learning the second pathway through the maze.

Recall test. The Transfer Group could have surpassed the Control Group in learning the second pathway only if it had learned the specific locations of the

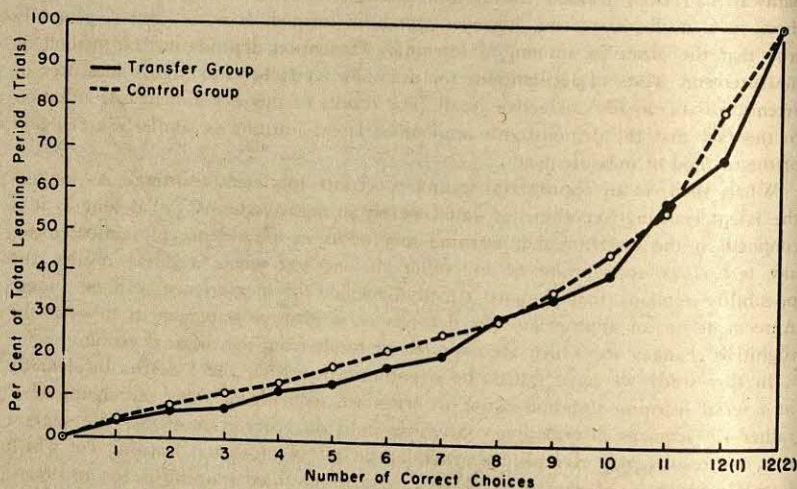


FIG. 1. VINCENT CURVES FOR ACQUISITION OF SECOND PATHWAY THROUGH THE MAZE

The point 12(1) refers to the first errorless trial; 12(2) is the criterion of two successive errorless trials.

irrelevant-incentive stimuli during the original training. The performance of the Recall Group provides us with an estimate of the amount of incidental memory for irrelevant-incentive stimuli regardless of location. The average number of such stimuli correctly reproduced by the Recall Group is 5.75, with an *SD* of 1.48.⁵ This average number is, of course, quite significantly different from zero ($t = 12.78$; $P = 0.001$). It is highly improbable that the recall score is seriously inflated by

⁴ Both the Transfer Group and the Control Group perform at better than chance level on the first trial of the second pathway. This result is probably due to the fact that they avoided the positions marking the first pathway which they knew to be no longer correct.

⁵ There is no correlation between the number of items recalled and the speed of original learning. The rank-order correlation of recall scores with trials to criterion is -0.08 ; with errors, 0.04 .

successful guesses; guessing would result in the reproduction of incorrect as well as correct words, yet there was not a single incorrect word among the items recalled. When performance on the critical test does not depend on memory for the specific location of the items, there is evidence for a considerable amount of latent learning.

DISCUSSION

Whether or not we find evidence for latent learning depends on the testing procedure we use. In this experiment, a test of transfer failed to yield any evidence whatsoever for latent learning of a pathway through a mental maze. A recall test, on the other hand, showed that almost half of the stimuli marking the irrelevant pathway had been learned during the original training. The specific location of these cues in the maze had, however, not been learned. It is, of course, generally true that the observed amount of learning or retention depends on the method of measurement. Tests of recognition, for example, typically yield higher measures of retention than do tests of active recall. The results of this experiment call attention to the fact that the demonstrable amount of latent learning is similarly a function of the method of measurement.

What, then, is an appropriate testing procedure for latent learning? As long as the latent learning hypothesis is stated merely in negative terms, *i.e.* as long as it is confined to the assertion that learning may occur in the absence of reinforcement, any test is as appropriate as any other. If one test yields negative results, the possibility remains that a more sensitive method of measurement can be found. Agreement on an appropriate test depends on a positive statement as to what the cognitive changes are which are expected to result from the original training.

In this study we have tested the hypothesis that what the *S* learns incidentally in a serial learning situation is not an organized sequence of signs and significates;⁶ rather he acquires information concerning individual irrelevant items. We suggest that our results may exemplify a general principle of incidental learning (of which latent learning is, of course, a special case): if rewarded training is for an organized sequence of responses to a series of stimuli, there will be incidental learning of the presence, but not of the serial arrangement, of the irrelevant stimuli. We would hope that this hypothesis will prove testable with animal *Ss* in the course of future experiments on latent learning. Whether or not this particular hypothesis proves useful, we wish to emphasize the need for positive hypotheses concerning the nature of latent learning.

SUMMARY

This experiment is concerned with the question of what it is the unrewarded *S* learns in an experiment on latent learning in which he is exposed to irrelevant-incentive stimuli. The results demonstrate that *S* may acquire knowledge of the presence of irrelevant incentives without learning their specific location. Whether or not evidence for latent learning is found depends on the testing procedure used. The need for the development of positive hypotheses concerning the nature of latent learning is stressed.

⁶ E. C. Tolman, *Purposive behavior in animals and men*, 1932, 135 f.

THE EFFECT OF VITAMIN-B DEFICIENCY ON THE HOARDING BEHAVIOR OF RATS

By NATHAN B. GROSS and VICTOR H. COHN, JR.,
Lehigh University

It has been demonstrated that when animals have been deprived of a specific nutritive material some of them, given a free selection of various types of food, will select those foods containing the needed material in preference to those which do not.¹ It seems a reasonable hypothesis that if some animals can select the material that they require, they might also *hoard* food containing it. Stellar attempted to study this question by injecting a variety of substances such as insulin and thyroid extract into rats,² and Bindra by depriving rats of various food substances including fat-soluble vitamins.³ Neither of these experiments yielded positive results, *i.e.* hoarding was not increased. The present experiment was designed to determine whether animals forced to live on a vitamin-B deficient diet for a prolonged period would hoard pellets containing a normal diet in preference to those containing the deficient diet.

Hoarding apparatus. The hoarding apparatus was divided into six alleys, each 40 in. long and 8 in. wide, built on a large piece of plywood. At one end of each alley was the animal's cage and at the other end a food bin. The food bin (7 in. deep) fit snugly into the alley and was divided into two equal parts by a partition. The back of the bin was sealed to prevent escape of the animals. Screening covered the alleys and food bins, permitting easy observation of the animals.

Food materials. All pellets were made with the same apparatus and were flat disks of approximately the same size ($\frac{3}{4}$ in. in diameter and $\frac{1}{8}$ in. thick). Pellets containing a normal diet were made by grinding Purina Dog Chow Checkers and adding milk-sugar as a binder and water. The average weight of these pellets was 1.86 gm. The pellets containing the vitamin-B deficient diet were made by mixing a standard vitamin-B deficient diet (68% sucrose, 18% vitamin test casein, 10% vegetable oil and 4% salt mixture U.S.P. XIV) with milk-sugar and water. The pellets averaged 2.56 gm.

Adaptation. The rats were fed exclusively on Purina Rabbit Chow from the

* Accepted for publication December 18, 1952. These experiments were made possible by a grant from the Institute of Research of Lehigh University.

¹ C. P. Richter, The internal environment and behavior. *Amer. J. Psychiat.*, 97, 1941, 878-893; P. T. Young, The experimental analysis of appetite, *Psychol. Bull.*, 38, 1941, 129-164.

² Eliot Stellar, The effect of epinephrine, insulin, and glucose upon hoarding in rats, *J. Comp. Psychol.*, 36, 1943, 21-31; The effects of experimental alterations of metabolism on the hoarding behavior of rats, *J. Comp. Physiol. Psychol.*, 44, 1951, 290-299.

³ Dalbir Bindra, Hoarding behavior of rats: Nutritional and psychological factors, Unpublished Doctoral dissertation, Harvard University, 1947.

time they were weaned until they were used in the experiment. They were then placed in their cages in front of the alleys for a period of three days so that they might adjust to the new situation. During this time they were provided with an excess of rabbit chow in the late afternoon and at night, but the food was removed in the morning, and no food was provided for about seven hours during the day. After the three-day adaptation period, each animal was permitted an opportunity to explore the alley and empty food bins.

Control. For each control test, normal pellets were provided in one bin and B-deficient pellets in the other, and the animal was given an opportunity to hoard for 30 min. The animal was permitted to eat from the pellets it hoarded. The type of pellet which each animal brought back to the cage on each trip was recorded. At the termination of the hoarding period, the pellets were removed and rabbit chow provided. Throughout the experiment the positions of the normal and deficient pellets were reversed after each hoarding period. The control tests lasted six or seven days. At the end of this period those animals which had hoarded less than about 50% normal pellets were chosen for the B-deficient experiment.

Deprivation period. After the control tests had been completed, the 17 animals chosen for the experiment were placed on the vitamin-B deficient diet for a period of 38-46 days. During this time an excess of the deficient food was provided. At the end of the deprivation period the animals were again tested daily for 6-7 days. The experiment terminated here for five of the animals.

Following the hoarding tests at the end of the deprivation period, 12 of the animals were then returned to the rabbit chow diet. In addition, a container of brewer's yeast powder was placed in the cage to hasten recovery. After the animals had been returned to the normal diet for 46-53 days, the 12 animals were retested.

Deficiency symptoms. Symptoms of vitamin-B deficiency were quite evident at the end of the deprivation period. The weights of the experimental animals were only one-third to one-half those of the litter mates not used in the deprivation experiment. All animals were undersized and had lost considerable hair. A few suffered from muscle spasms. When the animals were returned to the normal diet (rabbit chow and irradiated yeast) they showed marked gains in weight. Within a matter of days most of them weighed as much as their control litter mates. This rapid recovery of weight is particularly characteristic of animals who have been placed on a vitamin-B deficient diet and then returned to a fortified normal diet. It might also be added that the hair of the animals became more normal and the muscle spasms disappeared.

Statistical treatment of the data. It was judged best to treat the data in terms of the dichotomous proportions (normal vs. deficient pellets) and to determine the significance of the difference between these proportions. For each part of the experiment the percentage of normal pellets was determined for each animal. Percentages were then transformed to *thetas* and significance determined for the differences between *thetas*.⁴ This procedure was adopted in view of the skewed distribution of hoarding scores.⁵

Significance was determined for the following dichotomies: (1) the difference

⁴ P. O. Johnson, *Statistical Methods in Research*, 1949, 164-165.

⁵ C. T. Morgan, The statistical treatment of hoarding data, *J. Comp. Psychol.*, 38, 1945, 247-256.

between the percentages of normal pellets hoarded during the control period and at the end of the deprivation period; (2) the difference between the percentages of normal pellets hoarded during the control period and after the animal had been returned to the normal diet; (3) the difference between the percentages of normal pellets hoarded after the deprivation period and at the end of the time the animal was returned to the normal diet. In the ensuing discussion a difference will be considered as statistically significant if P is beyond the .001-level.

Results. The data obtained from 12 animals, selected from four groups of litter mates, are presented in Fig. 1. For each part of the experiment the total number of pellets hoarded (TL) and the number of normal pellets hoarded (NL) is presented.

In the hoarding tests following the deprivation period (black bars) all twelve animals hoarded significantly greater percentages of normal pellets than they had during the control tests. Six animals hoarded 95% or more normal pellets. Five of the remaining animals hoarded 75-85% normal pellets.

After the animals had been returned to the normal diet for periods of 46-53 days, all 12 of them hoarded significantly fewer normal pellets than they had at the end of the deprivation period. In general, their eating behavior tended to return to the level of the control period. Three animals hoarded about the same percentages as in the control period, two hoarded significantly less, and seven hoarded significantly more than during the control period (but again significantly less than at the end of the deprivation period).

The other five rats (four white and one hooded) were also tested before and after having been fed the deprivation diet for 38-39 days. All five animals hoarded significantly more normal pellets after the deprivation period. Because the animals were not returned to the normal diet, their results are not tabulated.

After the last hoarding test following the deprivation period each of the 17 animals was given a special mixed-pellet test. For this special test the partition was removed from the food bin and the normal pellets and deficient pellets were intermingled. Of the 17 animals, 16 hoarded significantly greater percentages of normal pellets than they had during the control period indicating that the preference was retained even when the pellets were mixed.

The three animals of Group I showed an unusual ability to discriminate between the normal and deficient pellets. During the control period the three animals of this group hoarded 56-65% deficient pellets. During the tests following the deprivation period Animal No. 1 hoarded no deficient pellets, Animal No. 2 hoarded one deficient pellet on each of the first three days, and Animal No. 3 hoarded two deficient pellets, both on the first day. Even during the special mixed-pellet test, the three animals of Group I hoarded only 4-6% deficient pellets. Hence it is notable that the animals of Group I, after having been returned to the normal diet for 53 days, hoarded approximately the same percentages of normal pellets as in the control period.

As a special control one animal was fed rabbit chow throughout the experiment. During the control period the animal hoarded 40% normal pellets. Then it was fed rabbit chow for forty days and retested. In the second test the percentage of normal pellets hoarded dropped to 10%. In this control, then, the percentage of normal pellets hoarded decreased rather than increased with the passing time.

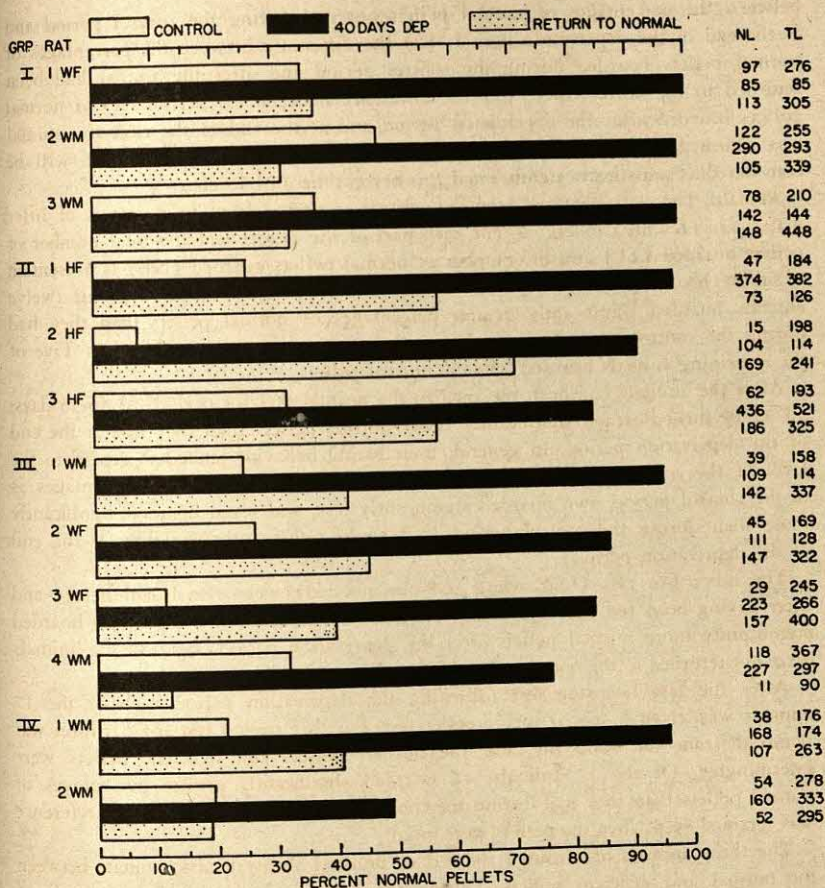


FIG. 1. THE EFFECT OF DEPRIVATION ON HOARDING

Percentage of normal pellets hoarded during the control period, after deprivation, and again after return to normal diet. *W*, white; *H*, hooded; *F*, female; *M*, male; *NL*, number of normal pellets; *TL*, total number of pellets hoarded. Information pertaining to nature and treatment of the *Ss* is here summarized:

Type	Group			
	I	II	III	IV
Number <i>F</i>	1	3	2	0
<i>M</i>	2	0	2	2
Age (in days) when placed on deficient diet	85	67	43	66
Days on deficient diet	41	46	39	45
Days on normal diet after deprivation period	53	46	47	51

The results of this experiment are in agreement with those obtained from the free-selection technique.⁶ A direct comparison with the studies of Stellar and of Bindra is not possible. The technique used in the present study differs from that used by Stellar, since natural depletion was accomplished and the deprivation was not a temporary condition induced by injection of a particular substance. The deprivation period in the present study far exceeded that employed by Bindra so that the difference between his results and those here reported might be attributed to the difference in length of the deprivation period employed.

Conclusion. Rats maintained on a vitamin-B deficient diet and tested in a hoarding situation show an increase preference for normal as compared with B-deficient pellets. The preference disappears when the animals have been returned to a normal diet.

⁶L. J. Harris, F. J. Hargreaves, and A. Ward, Appetite and choice of diet, *Proc. Royal Soc., Series B*, 113, 1933, 161-190.

THE RÔLE OF SET IN SUCCESSIVE DISCRIMINATION

By M. E. BITTERMAN and JAMES V. MCCONNELL, University of Texas

The relative difficulty of simultaneous and successive discrimination has been found to depend on the similarity of the stimuli to be discriminated¹ and on the nature of the apparatus employed.² The experiment here reported was designed to examine the effect of previous experience with the two types of problem on their relative difficulty. The importance of this variable was suggested by the incidental results of a recent investigation concerned with quite another question: One group of rats was trained on a simultaneous discrimination and a second group on a successive discrimination between horizontal and vertical stripes in a conventional jumping apparatus. Then the first group was trained on a simultaneous discrimination and the second group on a successive discrimination between two circles of different diameter. The performance of the simultaneous group was superior to that of the successive group on the initial problem, but the performance of the two groups on the subsequent problem did not differ significantly.³ These results suggest but do not unequivocally demonstrate the effect of previous experience on the relative difficulty of simultaneous and successive discrimination, since the difference in experience was confounded with a difference in the stimuli to be discriminated. The present experiment was designed to separate the two variables.

Subjects. Twenty-six naïve Albino rats bred in the laboratory were studied. They ranged in age from 4-5 mo. at the beginning of the experiment. The animals were paired for age and sex to form two matched groups (Simultaneous and Successive) of 13 animals each.

Apparatus. A conventional two-window jumping apparatus was employed. Behind the windows was a feeding platform to which the animal gained access following each correct response, and below them was a cloth net into which the animal fell after each incorrect response. The jumping platform was covered with a grid through which mild shock could be administered to break resistance to jumping. The entire apparatus was painted flat black.

Preliminary training. Throughout the experiment the animals were maintained on a 24-hr. feeding schedule. After an initial period of handling, during which they were fed wet mash on the feeding platform, the animals were taught to jump gradually increasing distances (to a maximum of 9 in.) from the jumping platform, first to open windows and then to unfastened mid-gray cards. Manual guidance was used from the beginning to provide equal experience with both windows.

* Accepted for publication January 19, 1953.

¹E. F. MacCaslin, Successive and simultaneous discrimination as a function of stimulus-similarity, this JOURNAL (in press).

²M. E. Bitterman and Jerome Wodinsky, Simultaneous and successive discrimination, *Psychol. Rev.*, 60, 1953, 371-376.

³Bitterman, A. D. Calvin, and C. B. Elam, Perceptual differentiation in the course of nondifferential reinforcement, *J. Compar. Physiol. Psychol.*, 46, 1953, 393-397.

Experimental training. Each group was trained on two discriminative problems in sequence—the Simultaneous Group on two simultaneous problems and the Successive Group on two successive problems. For 7 of the 13 pairs of animals, Problem I required a discrimination between black-and-white horizontally and vertically striped cards (the width of each stripe being 0.5 in.), and Problem II required a discrimination between two white circles of different size on black backgrounds (the diameter of the large circle being 4 in. and the diameter of the small circle 1 in.). For the remaining 6 pairs of animals, Problem I required a discrimination between the circles and Problem II required a discrimination between the horizontal and vertical stripes.

In the simultaneous problems, the two cards to be discriminated were paired on each trial; one of the cards was consistently reinforced and the other consistently punished. In the successive problems, the two cards presented on each trial were identical; the animal was reinforced for jumping to the right-hand member of one pair and the left-hand member of the other. Of the 13 animals constituting the Simultaneous Group, 3 were trained to jump to horizontal stripes and to large circle,⁴ 3 to horizontal stripes and small circle, 3 to vertical stripes and large circle, and 4 to vertical stripes and small circle. Of the 13 animals in the Successive Group, 3 were trained to jump left to horizontal stripes and to large circles,⁵ 3 to horizontal stripes and to small circles, 3 to vertical stripes and to large circles, and 4 to vertical stripes and to small circles.

Each animal was given 10 trials per day by the method of correction. Each trial consisted of a series of one or more jumps to a given pair of cards which terminated with a correct response. A maximum of three errors could be made on each trial, following which the animal was guided in the correct direction. The criterion of learning, both for Problem I and for Problem II, was two successive errorless days. The Simultaneous and Successive Groups were equated for amount of experience on the first problem by the following procedure: The first member of each pair to reach the criterion was overtrained on Problem I until the second member of the pair reached the criterion; then both animals were shifted to Problem II.

Results. In Fig. 1 the course of learning is plotted in terms of mean initial errors, and in Table I the performance of the two groups is summarized in terms of mean initial errors and mean total errors on each problem. The performance of the Simultaneous Group on Problem I was superior to that of the Successive Group. Differences in initial and in total errors, evaluated by Wilcoxon's non-parametric test for paired replicates,⁶ were significant beyond the 1-% level of confidence. On

⁴ Specifically, these animals were rewarded for jumping to the horizontally striped card, punished for jumping to the vertically striped card, rewarded for jumping to the large circle, and punished for jumping to the small circle. The phrases which describe the training of the other three Simultaneous sub-groups are to be interpreted in similar fashion.

⁵ Specifically, these animals were rewarded for jumping to the left window and punished for jumping to the right window when either two horizontally striped cards or two large circles were presented; they were rewarded for jumping to the right window and punished for jumping to the left window when either two vertically striped cards or two small circles were presented. The phrases which describe the training of the other three Successive sub-groups are to be interpreted in similar fashion.

⁶ Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures*, American Cyanamid Company, Stamford, Conn., 1949, 1-16.

Problem II, however, no significant difference between the groups could be found. There was a marked and significant improvement in the performance of the Successive Group from Problem I to Problem II, while the improvement of the Simultaneous Group was statistically insignificant.

In a recent paper on the relative difficulty of simultaneous and successive dis-

TABLE I
PERFORMANCE ON THE TWO SIMULTANEOUS AND SUCCESSIVE PROBLEMS

	Mean initial errors			Mean total errors		
	simulta- neous	successive	diff.	simulta- neous	successive	diff.
Problem I	16.3	27.2	10.9*	39.3	65.1	25.8*
Problem II	12.3	13.0	0.7	25.3	24.8	-0.5
Diff.	4.0	14.2*		14.0	40.3*	

* Significant beyond the 1% level by Wilcoxon's test for paired replicates.

crimination as a function of the kind of apparatus employed, Bitterman and Wodinsky provided an interpretation which applies equally well to the data of the present experiment.⁷ Let us assume (1) that two quite distinct sets may operate in the conventional jumping apparatus which was employed—the perception of the animal may be organized (a) in terms of the difference between the two pairs of cards as *pairs*, or (b) in terms of the difference between the two cards of each

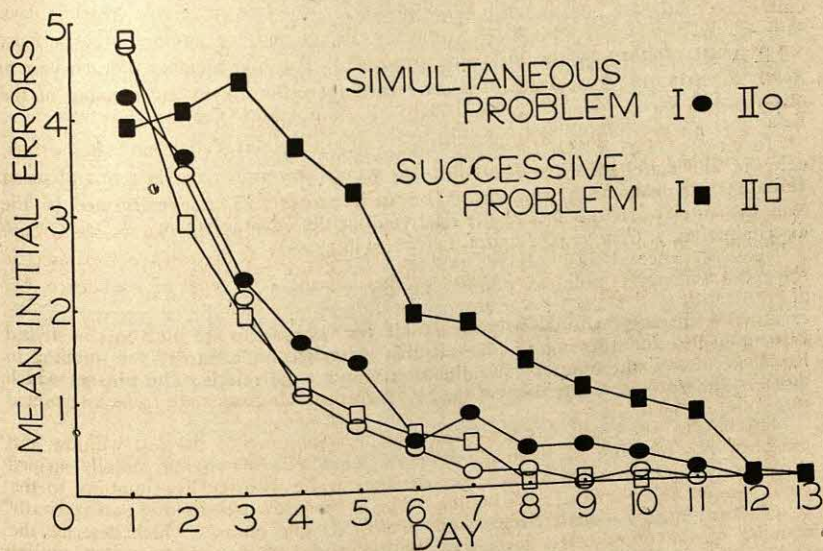


FIG. 1. THE COURSE OF LEARNING IN THE TWO GROUPS

⁷ Bitterman and Wodinsky, *op. cit.*, 371-376.

pair.⁸ In the first case we may speak of a *between-pairs* organization or a *configurational* set, while in the second case we may speak of a *within-pairs* organization or a set for *components*. We further assume (2) that the solution of a simultaneous problem is facilitated by a set for components and retarded by a configurational set, and that the solution of a successive problem is facilitated by a configurational set and retarded by a set for components. In the case of a simple black-white discrimination, for example, the two card-pairs of the simultaneous problem (black-white and white-black) are quite similar as pairs, while the difference between the members of each pair is quite striking. For the corresponding successive problem, however, the two pairs (black-black and white-white) differ markedly, while no solution in terms of components can develop except in relation to a spatial conditionality. Finally, we assume (3) that the conventional jumping apparatus, in which the animal must directly approach (jump at) one card of each pair, promotes the functional isolation of the two members of the pair, *i.e.* establishes a set for components and retards configurational organization. From these assumptions it follows that the initial difficulty of successive discrimination should be greater than that of simultaneous discrimination when the conventional jumping apparatus is employed.⁹ Once the configurational set has been achieved, however, the difference between the two types of problem should tend to disappear;¹⁰ transfer in the successive group should be large relative to that in the simultaneous group, a deduction which is confirmed by the results of the present experiment.

Summary. One group of rats (trained in the conventional jumping apparatus) learned two simultaneous discriminations in sequence while a second group learned two successive discriminations in sequence. The first successive problem was more difficult than the first simultaneous, but the second problems were equal in difficulty. The results are interpreted in terms of the transfer of set.

⁸ P. Weise and Bitterman, Response-selection in discriminative learning, *Psychol. Rev.*, 58, 1951, 185-195; D. C. Teas and Bitterman, Perceptual organization in the rat, *Psychol. Rev.*, 69, 1952, 130-140; J. R. Turbeville, Calvin, and Bitterman, Relational and configurational learning in the rat, this JOURNAL, 65, 1952, 424-433.

⁹ In an apparatus which facilitates configurational organization, the initial difficulty of the simultaneous problem should be greater than that of the successive (Bitterman and Wodinsky, *op. cit.*, 371-376; Wodinsky, M. A. Varley, and Bitterman, Situational determinants of the relative difficulty of simultaneous and successive discrimination, *J. Compar. & Physiol. Psychol.* (in press).

¹⁰ This statement holds only for problems involving relatively dissimilar stimuli, since the successive problem permits no direct comparison of the stimuli to be discriminated, and since in the absence of opportunity for direct comparison discrimination between similar stimuli is all but impossible (E. L. Saldanha and Bitterman, Relational learning in the rat, this JOURNAL, 64, 1951, 37-53). MacCaslin has demonstrated the effect of stimulus-similarity on the relative difficulty of simultaneous and successive problems (*op. cit.*).

DISCRIMINATION FOLLOWING VARYING AMOUNTS OF NONDIFFERENTIAL REINFORCEMENT

By M. E. BITTERMAN and CLAUDE B. ELAM, University of Texas

The data of a recent experiment by Bitterman, Calvin, and Elam led to the conclusion that nondifferential reinforcement at two points on an afferent dimension may facilitate subsequent discrimination between the points.¹ For example, two of the six groups of rats studied in the experiment were given 12 days of training in a conventional two-window jumping apparatus during which time all responses were reinforced; one group jumped to two identical mid-gray cards and the second to horizontally and vertically striped cards presented in simultaneous configurations (horizontal-vertical and vertical-horizontal). In the second stage of the experiment, both groups were trained on a simultaneous horizontal-vertical discrimination, and the performance of the second group was found to be superior to that of the first. This outcome suggested that a certain amount of perceptual differentiation had occurred in the course of nondifferentially reinforced experience with the striped cards. The present experiment was designed to study the effect upon subsequent discrimination of varying amounts of nondifferential reinforcement.

Subjects. Eighty naïve Albino rats, ranging in age from 3-4 mo., were studied. At the beginning of the experiment they were randomly divided into eight equal groups.

Apparatus. The animals were trained in a two-window jumping apparatus. Behind the windows was a feeding platform, and below them was a cloth net into which the animals fell after each incorrect response during discriminative training. The entire apparatus was painted flat black. The stimulus-cards were of three different kinds: (1) homogeneous mid-gray; (2) black-and-white horizontal stripes; and (3) black-and-white vertical stripes. The width of each stripe on each of the striped cards was 0.5 in.

Preliminary training. The animals were maintained on a 24-hr. feeding schedule throughout the experiment. After being handled and fed on the feeding platform for several days, they were taught to jump gradually increasing distances (up to a maximum of 9 in.) from the jumping platform to each of the open windows. Manual guidance was employed to provide equal experience with both windows, and each jump was rewarded with a nibble of wet mash.

Nondifferential reinforcement. In the first stage of the experiment, all groups were nondifferentially reinforced. Four of the groups (S-3, S-6, S-9, and S-12) jumped to the horizontally and vertically striped cards arranged in simultaneous configurations. The amount of experience given these four groups varied systemati-

* Accepted for publication April 15, 1953. We are indebted to Beverley Brand and E. D. Looney for assistance in the collection of data.

¹ M. E. Bitterman, A. D. Calvin, and C. B. Elam, Perceptual differentiation in the course of nondifferential reinforcement, *J. Compar. & Physiol. Psychol.*, 46, 1953, 393-397.

cally, the number of days of nondifferential reinforcement being indicated by the number used to designate each group (3, 6, 9, or 12). Each daily session consisted of 12 trials, 6 to each of the two lateral arrangements of the stimulus-cards. Of the 6 trials with each configuration, 3 were free and 3 were forced, and as a result each animal jumped equally often to each of the striped cards in each of the two windows. Each jump was reinforced with a nibble of wet mash. The remaining four groups of animals (G-3, G-6, G-9, and G-12) were treated exactly as were the correspondingly numbered S-groups, except that a pair of identical mid-gray cards was substituted for the striped cards. Of the 12 daily trials given these animals, 6 were free and 6 were forced, with the result that each animal jumped equally often to the two windows.

Discriminative training. In the second stage of the experiment (which began for each group on the day following the completion of the scheduled period of non-

TABLE I
ANALYSIS OF VARIANCE*

Source	df	Mean square	F	Sig.
Days	3	111	3.26	5%
Conditions	1	208	6.12	5%
Tests	1	1853	54.50	1%
Days×conditions	3	48	1.41	
Days×tests	3	89	2.62	
Conditions×tests	1	90	2.65	
Days×tests×conditions	3	18	0.53	
Within groups	64	34		
Total	79			

* The data analyzed are initial errors on the horizontal-vertical discrimination. *Days* refers to the number of days of previous nondifferential reinforcement, *conditions* to the use of gray as opposed to striped cards during the period of nondifferential reinforcement, and *tests* to the reinforcement of the horizontal as opposed to the vertical stripes in the discriminative problem.

differential reinforcement), all animals were trained on a simultaneous horizontal-vertical discrimination. For half the animals of each group (selected at random) horizontal stripes were reinforced, while for the remaining animals the vertical stripes were reinforced. Eight trials per day were given by the correctional method. Each trial consisted of a series of jumps terminating in a correct response which was either free or forced; the animal was allowed a maximum of three free jumps to a given arrangement of the cards, and after three errors it was manually guided in the correct direction. The two arrangements of the stimulus-cards were randomly varied from trial to trial. The criterion of learning was two successive errorless days.

Results. The results of the experiment are graphically summarized in Fig. 1, which shows mean initial errors on the horizontal-vertical discrimination as a function of number of days of nondifferential reinforcement on gray and striped cards. When total (initial plus repetitive) errors are plotted, much the same picture appears. The results of a statistical analysis of the data plotted in Fig. 1 are shown in Table I. Variance due to days of nondifferential reinforcement, to condition (nondifferential reinforcement on gray as opposed to striped cards), and to

test (reinforcement of horizontal as opposed to vertical stripes in the discriminative training) were statistically significant, but the interactions were not.

The advantage of the experimental design here employed becomes obvious when the results are compared with those of Bitterman, Calvin, and Elam. In both experiments, animals with 12 days of nondifferential reinforcement on the striped cards make significantly fewer errors on a subsequent horizontal-vertical discrimination than do animals with 12 days of nondifferential reinforcement on grays. Only when the amount of nondifferential reinforcement is varied, however, does it become apparent that the transfer-effect of nondifferential reinforcement (whether on stripes or grays) is predominantly negative.

The present data also serve to clarify the relation between the earlier results and

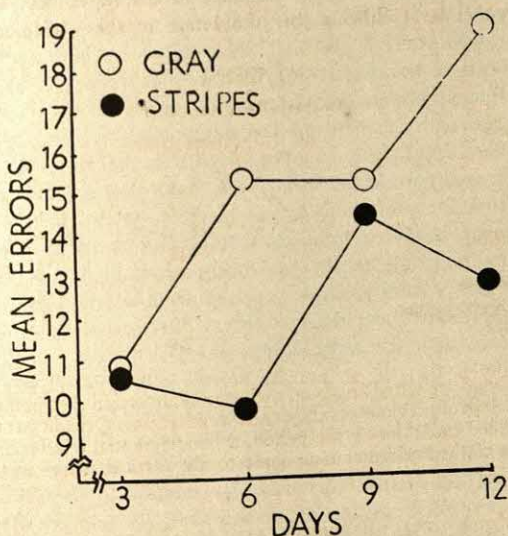


FIG. 1. MEAN INITIAL ERRORS ON THE HORIZONTAL-VERTICAL DISCRIMINATION AS A FUNCTION OF AMOUNT OF PREVIOUS NONDIFFERENTIAL REINFORCEMENT ON GRAY AND STRIPED CARDS

those obtained by Siegel several years previously.² Siegel gave rats two degrees of nondifferential reinforcement in a simple T-maze, following which the animals were trained to make a simple positional response. In the second stage of the experiment the performance of the group with the greater amount of nondifferential reinforcement was slightly but significantly inferior. This result, which suggested that perceptual differentiation (insofar as it is involved in the mastery of a simple T-maze) is retarded by nondifferentially reinforced experience, seemed contrary to the concept of perceptual learning and seemed to support a deduction by Hull which was based on the concept of a negatively accelerated growth of habit-strength

² P. S. Siegel, The rôle of absolute response strength in simple trial-and-error learning, *J. Exper. Psychol.*, 35, 1945, 199-205.

as a function of reinforcement.³ The results of Bitterman, Calvin, and Elam, however, clearly contradicted the Hullian deduction and implied that perceptual differentiation may occur in the absence of differential reinforcement. The present experiment confirms both sets of findings—the results for the S-groups alone being directly analogous to those obtained by Siegel.

Explanation of the data here reported appears to involve at least two effects. The general tendency for nondifferentially reinforced experience (whether with stripes or with grays) increasingly to retard performance on the subsequent horizontal-vertical discrimination suggests an attentional modification—a decrease in the functional significance of certain features of the situation which are relevant to the subsequent discriminative task—whether it be interpreted in terms of perceptual selectivity or of orienting response. The data of Walker, Knotter, and DeValois warrant the hypothesis that the effect may be related in part to the use of forced-choices during the first stage of the experiment.⁴ Whatever the source of the negative transfer, it does not seem to be related to changes in the strength of positional preferences. During the first 9 days of nondifferential reinforcement the strength of these preferences remained very much the same, both from day to day and from group to group; during the last three days the preferences were actually somewhat weakened in G-12 although they rose slightly in S-12.

The general superiority of the S-groups to the G-groups in the second stage of the experiment may be taken to mean that nondifferential reinforcement on the stripes contributes positively to the subsequent discrimination. It is possible that the horizontally and vertically striped cards are not at the outset maximally differentiated for the animal, and that perceptual differentiation occurs in the course of sheer experience with the cards despite the lack of differential reinforcement. Another possibility is that the stripes are perceptually differentiated to begin with, and that their presence during the first stage of the experiment merely serves to forestall the negative developments already considered. A third possibility is suggested by the significant variance due to test—by the fact that the performance of animals trained vertical-positive was markedly inferior to the performance of the animals trained horizontal-positive. If nondifferentially reinforced experience with the stripes tended partially to overcome an initial aversion to the vertical, the performance of the horizontal-positive group would not be appreciably impaired and the performance of the vertical-positive group would be markedly improved (relative to the performance of the corresponding G-groups).

Of these three possibilities, the third can quite definitely be rejected. While the S-groups did show a slight preference for the horizontal stripes in the first stage of the experiment (making an average of 2.7 out of 6 daily free jumps to the vertical), the magnitude of the preference was not systematically changed in the course of nondifferential reinforcement. Furthermore, on the first day of training in the second stage of the experiment the corresponding S-groups and G-groups showed closely similar preferences; mean responses to vertical were 2.5, 3.3, 2.8, and 3.3 for groups G-3, G-6, G-9, and G-12, and 2.9, 3.8, 2.5, and 3.1 for the corre-

³ C. L. Hull, *A Behavior System*, 1952, 21-29.

⁴ E. L. Walker, M. C. Knotter, and R. L. DeValois, Drive specificity and learning: The acquisition of a spatial response to food under conditions of water deprivation and food satiation, *J. Exper. Psychol.*, 40, 1950, 161-168.

spondingly numbered S-groups. Finally, the third interpretation does not account adequately for certain previous findings: (1) in the experiment by Bitterman, Calvin, and Elam, a group which was nondifferentially reinforced on the striped cards presented in successive configurations (horizontal-horizontal and vertical-vertical) was not superior to the group which was nondifferentially reinforced on grays, despite the fact that there was opportunity in the former group for equalization of stripe-preferences; (2) in the same experiment nondifferentially reinforced experience with the striped cards in successive configurations facilitated performance on a successive horizontal-vertical discrimination in which stripe-preferences were ruled out and in which corresponding configurational preferences did not exist; (3) in an experiment by Bitterman, Elam, and Wortz,⁵ extensive amounts of nondifferential reward and punishment did not significantly affect performance on a simultaneous horizontal-vertical discrimination when the nondifferentially reinforced experience with the stripes was given in the context of a problem to the solution of which the horizontal-vertical difference was irrelevant.

It is unfortunate that the data of the present experiment do not more adequately define the shapes of the curves plotted in Fig. 1. Clearly the two functions must take their origin from a common point, since with zero days of nondifferential reinforcement the two conditions are functionally identical ($S-0 \equiv G-0$). If, then, as the amount of nondifferential reinforcement increases, the performance of S-groups becomes superior to that of G-groups, the two curves must be expected to diverge, as indeed they appear to do despite the fact that the interaction of days $\times$ conditions does not reach statistical significance. The subsequent course of the curves can be determined only by an experimental examination of the effects of longer periods of nondifferential reinforcement.

SUMMARY

Rats were trained in the jumping apparatus to discriminate between horizontally and vertically striped cards after 3, 6, 9, or 12 days of nondifferential reinforcement on either the two striped cards or on a pair of identical mid-gray cards. Performance on the discriminative problem tended to deteriorate with increasing amounts of nondifferential reinforcement both on stripes and on grays, but the deterioration was more marked in the groups nondifferentially reinforced on grays. The results suggest a general retarding effect of nondifferential reinforcement which is compensated for in part by experience with the striped cards.

⁵Bitterman, Elam, and E. C. Wortz, Perceptual differentiation as a function of nondifferential reward and punishment, *J. Compar. & Physiol. Psychol.*, 46, 1953, 475-478.

LEARNING AS A FUNCTION OF WORD-FREQUENCY

By JOHN F. HALL, Pennsylvania State University

The experimental work of Howes and Solomon has indicated a high inverse relationship between the frequency with which a word appears in the language (as indicated by the Thorndike-Lorge word count¹) and the duration of time required for S to perceive it.² In an effort to provide a more exact statement of this relationship, Solomon and Postman controlled frequency more precisely by having Ss read and pronounce different nonsense words with frequencies ranging from 1 to 25³. The results indicated the same inverse relationship between the recognition thresholds and the frequency of prior usage.

If the assumption is made that there is a relationship between perceptual threshold and learning as measured by a recall score (that those words perceived most rapidly are easiest to learn), words which appear with a high frequency in the language should be learned more readily than those of low frequency. Aside from this theoretical consideration, the relationship found between learning and word-frequency should be of value in the calibration, construction, and utilization of word-lists which are so frequently used in rote-learning experiments.

It should be noted that an earlier study by Peters did not seem to support the existence of such a relationship.⁴ In one phase of a series of experiments, Peters constructed five different word lists, the words in each list being selected from the 2500 most frequently used words found in the *Teacher's Word Book of 10,000 Words*.⁵ List I contained words with frequencies of 1 to 500, List II contained words with frequencies of 501-1000, List III contained words with frequencies of 1000-1500, and so forth. The Ss were then presented with these lists in a counter-balanced order. Recall scores obtained after one presentation indicated no significant differences between lists.

The relatively restricted range of frequencies that Peters employed, in contrast to the range employed by Howes and Solomon, would seem to make a more extensive test of the relationship between word-frequency and recall desirable. The present experiment was designed to provide such a test.

Frequency of word-usage. A word-frequency count is obtained by selecting a given

* Accepted for publication April 16, 1953. This study was supported by a grant from the American Philosophical Society.

¹ E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944, 1-274.

² D. H. Howes and R. L. Solomon, Visual duration thresholds as a function of word probability, *J. Exper. Psychol.*, 41, 1951, 401-410.

³ Solomon and L. Postman, Frequency of usage as a determinant of recognition thresholds for words, *J. Exper. Psychol.*, 43, 1952, 195-201.

⁴ H. N. Peters, The relationship between familiarity of words and their memory value, this JOURNAL, 48, 1936, 572-584.

⁵ Thorndike, *The Teacher's Word Book of 10,000 Words*, 1921.

amount of written material and then tabulating the number of times that each word contained in that material appears. Probably the best source of such a word-count is *The Teacher's Word Book of 30,000 Words*. Here the authors report tabulations from four sources: (1) textbooks, readers, the Bible, and the English classics; (2) recent and popular magazines; (3) books recommended for boys and girls in grades 3 to 8; (4) a miscellany of juvenile and adult reading. Though counts for the four sources are reported separately, the authors point out that the most important and usually decisive index is the summary which combines all four. This summary was used as the guide in the selection of words used in the present study.

Materials. Four word-lists, each list containing 20 words, were used. *List A* contained words which had a frequency-count of 1 occurrence per million words; *List B* had a frequency-count of 10 occurrences per million words; *List C* had a frequency-count of 30 occurrences per million words;⁶ and *List D* had a frequency-count of 50 to 100 occurrences per million words. The four lists represented a range of more than the 10,000 most frequently used words.

The words used in each of the lists were obtained in the following manner: A table of random numbers was first consulted to provide a list of page-numbers with only numbers ranging from 1-208 (the number of pages in the Thorndike-Lorge text) being included. The first word, on the page indicated by the first random number, which met the required frequency and consisted of seven letters was noted. This last requirement was made in order to insure complete uniformity with respect to word-length. If the page did not have a word with the desired frequency and correct length, successive pages were examined until such a word was obtained. The serial positions of the words in each list were obtained by assigning each word a number and then consulting a table of random numbers to determine its position.

The words used in each list and their order of presentation were as follows:

List A. winsome, outcrop, saddler, foolery, bagpipe, timbale, muffler, disport, angling, languor, laconic, diurnal, cession, swaddle, interim, redoubt, emulate, missive, impiety, derange.

List B. astound, discard, pastime, forlorn, esquire, pitiful, luggage, implore, storage, haughty, freeman, lowland, moisten, ecstasy, chemist, mariner, rapture, valiant, ghastly, cleanse.

List C. upright, immense, revenge, chimney, earnest, contact, pasture, absence, majesty, scratch, confine, doorway, blanket, despite, missing, heavily, furnace, sparkle, violent, harness.

List D. forward, because, destroy, special, against, weather, another, example, problem, quarter, instead, believe, declare, outside, passage, written, silence, morning, teacher, attempt.

Procedure. Nine groups of 20 to 30 students each, recruited from introductory classes in psychology at Pennsylvania State University, served as Ss. In all, 227 Ss were used. By means of a slide-projector, each word from one of the lists was presented to a group for 5 sec., with approximately 120 sec. required to present the total list. Four such presentations were made with a 30-sec. rest-period between

⁶ The value for *List C* is not wholly correct. There was not a sufficient number of 7-letter words having a frequency count of 30 to construct a list of 20 words; consequently, it was decided to use words having frequency counts of 29 and 31. *List C* contains 10 words having a frequency count of 29, 7 with a frequency count of 30, and 3 with a frequency count of 31.

trials. After the last trial, paper was distributed and the Ss were given 5 min. to print as many of the presented words as they could recall.

Results. Performance was measured in terms of the number of words correctly written by each S. Misspellings or variants of the correct word were not counted. The results, which are summarized in Table I, confirm the hypothesis that, within

TABLE I
MEAN NUMBER OF WORDS RECALLED AS A FUNCTION OF WORD-FREQUENCY

List	Frequency	N	Mean	Diff.	t	Sig.
A	1	76	12.04			
B	10	52	13.31	1.27	2.02	5%
C	30	44	15.02	1.71	2.50	5%
D	50-100	55	15.04	0.02	0.02	—

limits, the more frequently a word appears in the language, the more readily it is learned. Indirectly, therefore, they also tend to support the existence of a relationship between perceptual threshold and learning. From a methodological point of view, these results suggest that the Thorndike-Lorge word-count can be used, within limits, to calibrate words which may be used in learning experiments. Interestingly enough, these findings do not conflict with the earlier study by Peters, who used lists selected from the 2500 most frequently used words, which according to Thorndike and Lorge, have a frequency-count of 30 and above. It will be noted that in the present study, the two lists constructed within this range of word-frequencies, did not show a difference in recall-scores.

Summary. Word-lists, containing words of a different frequency as indicated by the Thorndike-Lorge word-count, were presented to groups of college students, who were subsequently tested for recall. A significant positive relationship between word-frequency and recall was found.

AN ANCHORING EFFECT IN THE OPERANT RESPONDING OF RATS

By EDWARD J. GREEN, Harvard University

Skinner reported in 1938 that high levels of illumination suppressed operant response rate in the white rat.¹ The present study was aimed at exploring the effective range of suppression.

Procedure. A modified Skinner box made with bar and food magazine placed at one end was used. An illuminated milk glass panel $2\frac{1}{2} \times 3\frac{1}{2}$ in. was fixed to the cage directly above the bar and 7 in. above the floor. Water was always available to the animals, air was continuously changed by an exhaust fan, and white noise was introduced for masking. The animals were four Wistar albino rats, approximately 200 days old at the start of the experiment. They were deprived of food until they reached 80% ad libitum weight. They were then run on a 2-min. variable interval reinforcement schedule until their rates of response had stabilized. This initial training period and the 10-day runs after the rates had stabilized were conducted with a brightness of 26 millilamberts (mL.) measured at the milk glass panel.

During the next 30 days the rats were run 1 hr. a day with brightness changing once a minute, independently of both response and reinforcement. Twelve different brightness levels were used. Apparatus limitations permitted the use of only four levels during any one experimental period, so three ranges of four brightness levels each were used. These were randomized in order of presentation, and each rat was subjected to each range for ten 1-hr. sessions. Each brightness range overlapped the next in the following manner. The low range extended from 0.5 mL to 26 mL; the medium range from 24 to 376 mL; and the high from 149 to 1166 mL. For the sake of simplicity, the brightest level in each range is called Stimulus I, and the other levels II, III, and IV.

After this 30-day period, 10 more days were run under the original conditions of a constant brightness of 26 mL.

Results. Total responding between the three brightness ranges was not significantly different; responding within each range, however, exhibits the expected effect of increased illumination. The number of responses emitted by each rat to each stimulus within a range was ordered with respect to the order of stimulus-intensity, and a non-parametric test of significance was made. The orderings for the middle and high ranges were significant beyond the 0.001 level. The ordering for the low range was not significant.

The surprising fact of this experiment was the peculiar overlapping of response rates between the three brightness ranges. In Table I, one sees that responding to the

* Accepted for publication January 31, 1953. This study was directed by Dr. W. S. Verplanck.

¹B. F. Skinner, *The Behavior of Organisms: An Experimental Analysis*, 1938, 254-257.

lowest level of the middle range is much greater than responding to any of the brightness levels of the lower range. Almost the same is true of responding to Stimulus IV in the high intensity range as compared with both the low and the medium ranges: a greater number of responses was emitted only to Stimulus IV of the middle range.

It seemed that contrast effects arising from the serial position of a given stimulus intensity might be responsible for the phenomenon, so the number of responses made under Stimulus I was partialled into four categories: those occasions in which this intensity immediately followed Stimulus IV, III, II, and itself. A similar breakdown

TABLE I
MEAN NUMBER OF RESPONSES PER 15 MIN. EMITTED UNDER EACH
EXPERIMENTAL LEVEL OF BRIGHTNESS

Stimulus	Brightness range		
	Low	Middle	High
I	113	93	84
II	126	109	100
III	129	133	121
IV	136	150	147

was made for responses made to Stimulus IV. The records for two days under each of the three ranges were used for this analysis, since readable records were not available for every session. It was thought that such a sample would be sufficient for non-parametric statistical treatment. The same X^2 statistic was applied to these data as was used with the orderings mentioned above. For the Stimulus I breakdown, the orderings were not at all significant. For the Stimulus IV breakdown, p fell between 0.10 and 0.20. Certainly this is no more than a suggestion that serial position or contrast alone have produced this effect.

The study has demonstrated that the rate of operant responding in the white rat is inversely proportional to intensity of illumination, although the illumination is not contingent upon anything the rat does. It also seems that if a low rate of response is to be taken as a measure of aversiveness, then aversiveness must be defined in terms of the relative position of the stimulus with respect to others rather than by any absolute criterion. It is suggested that the observed phenomenon may be classed as an example of 'anchoring' in the operant behavior of the white rat.

APPARATUS

AN APPARATUS FOR STUDYING THE PERCEPTION OF LIGHT-MOVEMENTS

By J. H. ROHRER and E. L. HOFFMAN, Urban Life Research
Institute, Tulane University

The apparatus was designed to meet the following requirements: (1) a light-source should be moved at a given constant speed over a given distance; (2) the direction of movement should be readily variable; (3) the movement should be automatically repeatable, *i.e.* the light-source turned off, the lamp-housing returned to its point of origin, and the movement begun again; (4) the timing controls should make it possible to turn on and off a stationary light-source while attendant cues produced by movement of the lamp-housing are provided.

Description. The assembled apparatus is pictured in Fig. 1. The mechanism is mounted on a $\frac{1}{4}$ -in. steel slab, 5 in. wide and 18 in. long. The slab, in turn, is mounted on a pyramidal iron framework supported at its base by a welded foundation 16.5 in. wide and 18 in. deep. A steel disk, 7.5 in. in diameter, provides the surface for mounting the slab. Holes $\frac{1}{2}$ in. in diameter are center-drilled in the disk $\frac{1}{2}$ in. from its outer edge at 0° , 45° , 90° , 135° , 180° , 225° , 270° , and 315° . A spring-tensioned ball-bearing catch is mounted on the slab adjacent to the disk at a 3.5-in. radius from point of rotation of the slab, thus providing a way of locking the slab in any of the eight positions on the disk to which it may be manually rotated. The mechanical power-source utilized is a model-11 governed-speed reversible motor with a r.p.m. range of 500-7500, a product of the Lee Engineering & Manufacturing Company, Milwaukee. A 6-v. radio panel-lamp is used as the light-source. A wiring diagram of the apparatus is given in Fig. 2. The identifying code-letters are the same in both figures.

The lamp housing, made from a copper can 1.875 in. in diameter and

* The apparatus described was constructed for research under ONR Contract Nonr-475(01) with Tulane University and the Office of Naval Research, in cooperation with the Neuropsychiatry Branch, Bureau of Medicine and Surgery, U. S. Navy.

0.625 in. deep, is attached to a platform driven by the screw C. The front of this housing has drilled in it a $\frac{1}{4}$ -in. hole, located $\frac{1}{2}$ in. off center. The entire front of the housing is covered with a cap in which three apertures—1, 2, and 3 mm. in diameter—are drilled $\frac{1}{2}$ in. off-

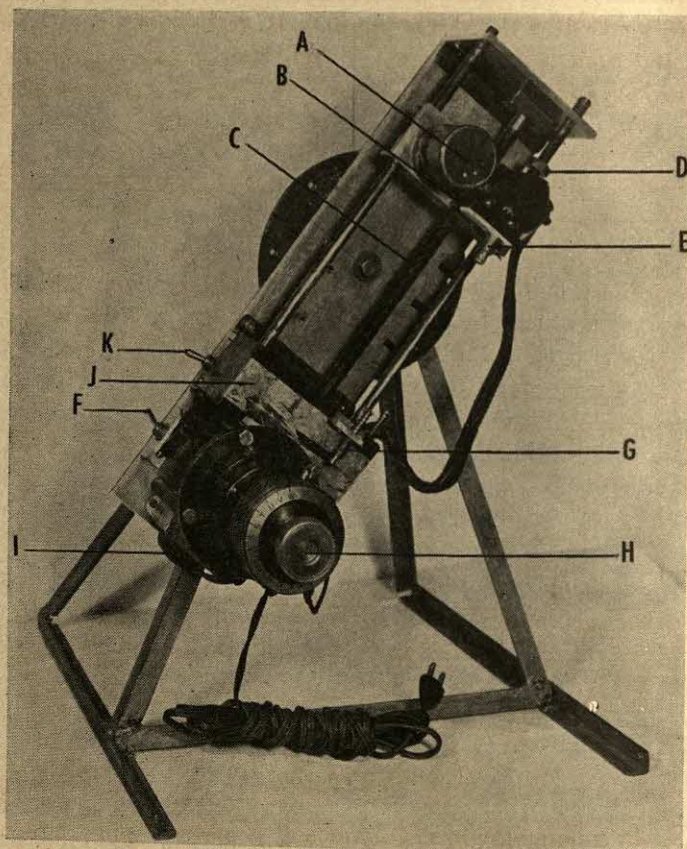


FIG. 1. PHOTOGRAPH OF THE APPARATUS
(For description see text.)

center. Any one of these apertures may be aligned with the $\frac{1}{4}$ -in. hole in the front of the light-housing by adjusting the position of the cap on the housing. One or more sheets of diffusing tissue may be placed between the cap and the housing to alter the intensity of the light. The circuit connecting this moving light with switches B and F and motor

windings *H* may be diverted with the use of jack plugs to a stationary light that can be used as a stimulus for the autokinetic phenomenon. The stationary light utilizes the timing features and attendant sound-cues of the movement-apparatus.

Operation. As the light-housing-platform (*LHP*) reaches bottom of screw *C*, switch *B* is closed, causing the relay-coil *J*, the 6-v. transformer *I*, and the lamp *A* to be energized. The relay-contacts so connect the motor-windings *H* as to turn the screw *C* to drive the *LHP* upward. At a

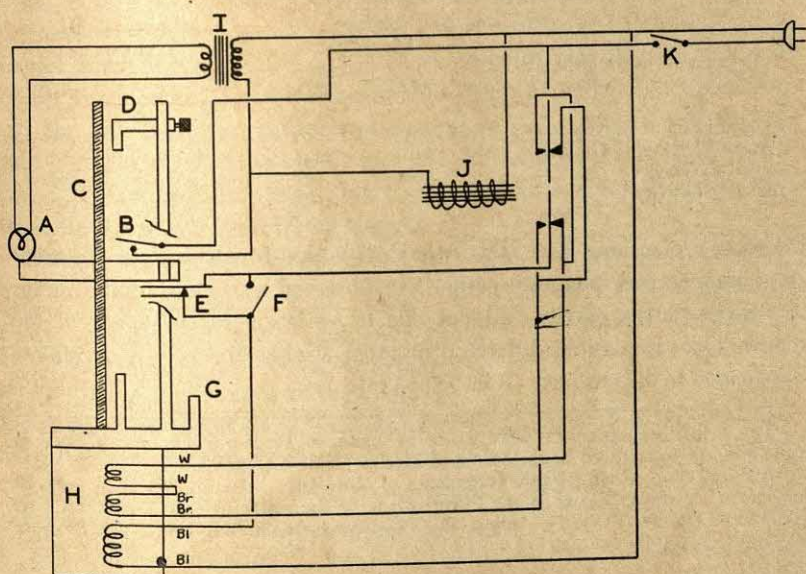


FIG. 2. SCHEMATIC WIRING DIAGRAM
(For description see text.)

point set by adjustable probe *D*, switch *B* is opened, lamp *A* is turned off, and relay-coil *J* is de-energized. The relay-contacts reverse the motor-windings *H*, and screw *C* reverses, driving *LHP* downward. If switch *F* is closed, the *LHP*, upon reaching bottom, will cause switch *B* to close and the operation will repeat. If switch *F* is open, contact *E* will be opened by a probe, opening a motor-winding and stopping the operation. Switch *K*, when closed, provides for the flow of a 110-v. alternating current from the power-source to the transformer and motor windings through switches *F* and *K*.

Constancy of the speed of movement. The reliability of the apparatus for producing a constant speed of movement is documented in Table 1.¹ Using Aubert's data for discriminable differences,² the rate of movement toward a gravity pull, at the slowest movement of which the apparatus is capable, would be discriminated as different from movement along

TABLE I

MEANS AND STANDARD DEVIATIONS OF TIME (IN MINUTES) FOR LIGHT MOVEMENT OVER AN 8-IN. DISTANCE AT THE MOST RAPID AND AT THE SLOWEST SPEED
(TEN MEASUREMENTS AT EACH POSITION)

Direction of Movement	Slowest Movement		Most Rapid Movement	
	M	SD	M	SD
Toward Pull of Gravity	.493	.007	.072	.001
Away from Pull of Gravity	.616	.013	.071	.001
90° R to Gravity	.590	.008	.071	.002
90° L to Gravity	.596	.009	.070	.001

the other dimensions. The SDs of the distribution for the various directions are homogeneous, suggesting that the observed variation in the means is attributable to gravity and not due to variation in the speed of the motor. This discriminable difference disappears when the rate of movement is increased to 0.9 cm./sec. (8 in. / 0.33 min.).

¹ The reliability measures were made by using a Type S-6 Timer calibrated in .001/min., manufactured by the Standard Electric Time Company.

² Aubert reports that reliable judgments of difference between two basic rates of movement can be made when the difference is of the order of 1 to 2 mm. arc./sec. (0.05-0.10 cm./sec.). See H. Aubert, *Die Bewegungsempfindung*. *Arch. ges. Physiol.*, 39, 1886, 347-370.

AN OLFACTOMETER PERMITTING STIMULUS SPECIFICATION IN MOLAR TERMS

By F. NOWELL JONES, University of California, Los Angeles

Thresholds for chemoreceptors—olfactory or gustatory—are best given in terms of the molecular concentration of the stimulating substance in the medium used to convey it to the receptor-surface. This implies that control must be exercised over all other stimulus-variables—in the case of olfaction, duration of the stimulus-blast, pressure, volume, temperature, and humidity. Some of the previous attempts to accomplish this control have been reviewed by Wenzel,¹ and the difficulty encountered when an aspect of the stimulus other than concentration is used to define the threshold has been demonstrated elsewhere.²

Difficulties of design center around two main problems, first, accuracy in specifying stimulus-concentration, and second, cost. Having a quantity of pure air with a known stimulus-concentration, it is relatively easy to devise a means of delivering a controlled blast of it to S's nose, although doing so inexpensively is not too easy. The solution to these simultaneous problems to be presented below is not the only one possible, but it has proved relatively simple and convenient.

A diagram of the apparatus is shown in Fig. 1. The odorous stimuli are contained in a series of evaporation-bottles. Each unit (*A*) really consists of two bottles containing identical materials, so that as odorous air is drawn off, dilution progresses slowly, and each bottle has a glass-wool 'wick' to hasten evaporation. In each absorption-unit, the odorous material is dissolved in mineral oil, so that the odorous air is not too far above the expected threshold-concentration. When necessary to cover the range of individual differences, two or three different dilutions may be prepared. Since there is no bubbling of air, there is no chance of aerosol formation.

* The work reported herein was performed under Contract No. N6onr-27515, Project NR140-008 between the University of California and the Office of Naval Research.

¹ Bernice M. Wenzel, Techniques in olfactometry: A critical review of the last one hundred years. *Psychol. Bull.*, 45, 1948, 231-247.

² F. N. Jones, A test of the Elsberg technique of olfactometry, this JOURNAL, 66, 1953, 81-85; J. A. Whittenburg, The effects of general activity and extended diurnal variation on olfactory sensitivity. Technical Report No. 11, Project DA-49-007-MD-222(O.I, 19-52), Army Medical Research and Development Board.

For each odor a syringe (*B*) three-way stopcock (*C*) and nosepiece (*D*) are provided. These are clamped into holders on the main body of the apparatus, and connected to the absorption bottles by means of a short length of silicone-rubber tubing.³ All connections among parts are made with silicone rubber, and all rubber stoppers are covered with aluminum foil. The nosepieces are especially fabricated, and provide air-tight connec-

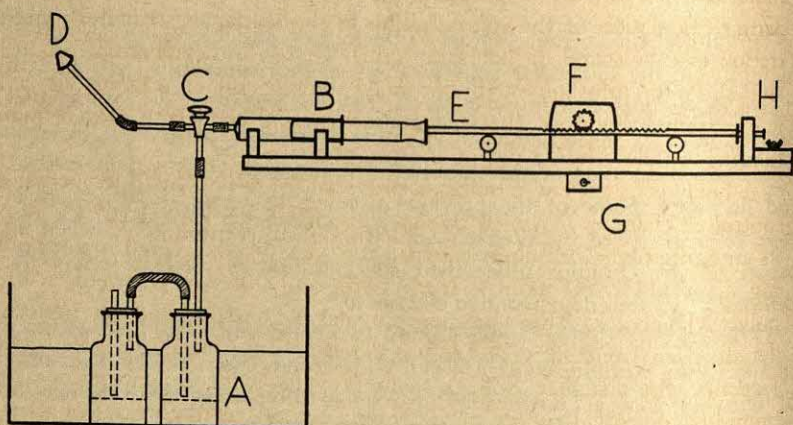


FIG. 1. DIAGRAM OF THE ESSENTIAL PORTIONS OF THE OLFACTOMETER

tions to the nostril. When clamped in position, the plunger of the syringe makes contact with a rack, (*E*) which is driven by a constant-speed motor, (*F*) controlled by switch, (*G*). A stop (*H*) and the length of the toothed portion of the rack determine respectively the point at which the movement of the rack begins and the point at which it terminates. Since the stop and the motor-mount are adjustable, the extent of movement, and therefore the amount of air forced from the syringe, may be readily varied. As presently geared, a blast of 40 cc. is delivered in 0.8 sec. These magnitudes are such as to minimize quantity of stimulus-blast as a determiner of threshold. The absorption-units are stored in a constant-temperature bath and are immersed in a constant-temperature bath during experimentation. The temperature of the baths is 25.9° C. with a maximum variation of $\pm 10.1^\circ$ C. Pure air is supplied continuously through a filter-chain consisting of silicia gel, activated charcoal, and silicia gel, in that order. This air is warmed to the same temperature as the stimulus bottles, and led to two funnels, not shown in the diagram, which serve respectively to provide a pure-air replacement

³ Obtained from the Connecticut Hard Rubber Company, New Haven.

for odorous air drawn off from the stimulus-bottles, and a source of pure air for the stimulus-mixture.

Thresholds are found by metering various amounts of odorous air into the syringe by proper adjustment of the stopcock, (C) and then filling the syringe with pure air. This provides a very fine control of concentration and, since a knowledge of the vapor-pressure of the odorous material, its molecular weight, the molecular weight of the oil, and the temperature, permit calculation of the concentration in the bottle, the stimulus-concentration may be stated in molecular terms. There is no evidence of progressive shifts in threshold with continued drawing off of odorous air, and it is assumed therefore, that evaporation is rapid enough to prevent more than negligible dilution during a series of trials.

A fairly constant environment is provided for *O* by a small temperature-controlled chamber in which he sits. This chamber is kept at 20° C., all of its air being filtered through charcoal. It is glass-lined, except for zinc holding strips and a few stainless steel inserts, and *O*'s chair is covered with a polyethylene sack. This chamber provides a constant, but not perfectly 'baseline' environment. Constancy is aided to a considerable degree by the fact that the chamber is in a room provided with a temperature- and humidity-controlled supply of air from a central ventilating system. Temperature variation in the chamber is very slight—too small to be measured with an ordinary laboratory thermometer.

The experimental routine for obtaining a single threshold is as follows: After all connections have been made, a small amount of odorous air is drawn into the syringe, and then the syringe is filled with pure air by retracting the plunger while the nosepiece is placed in the pure-air funnel. Now *O* places the nosepiece in his nostril, and the blast is delivered by switching on the motor. After every trial, the rack and plunger are withdrawn after lifting the motor a small amount on its hinged mounting, and while the nosepiece is in place in the pure-air funnel. Several purging strokes of the plunger are then given by hand, and another trial begun. In actual practice, events proceed very rapidly—as rapidly, indeed, as trials should be given, since adaptation must be avoided. After use by each *O*, the nosepieces are washed and sterilized with ultra-violet light.

Sample results from four *O*s are available. In order to test the accuracy of the apparatus, it was decided to use two dilutions of two substances. Thus, it is possible to decide if the dilution in mineral oil and the metering provided by the syringe yield comparable results. That is, even if the concentration of odorous molecules in the absorption-bottles differs, the thresholds in molecular terms for the same substance should remain the same.

The two substances used were methyl salicylate and benzene.⁴ There were unexpectedly large individual differences in sensitivity to methyl salicylate, and cross-comparison was not possible except for one inexperienced *O*. For benzene, however, comparison of two levels of dilution was made for all four *O*'s. These results, each threshold being an average of three ascending determinations, are given in Table I. Correspondence is quite close, considering the small number of trials. It should be pointed out that these thresholds are for recognition, which we have found to be easier to judge than are absolute thresholds. Shown also in Table I are the thresholds for methyl salicylate.

The method of calculating thresholds in terms of molar ratios may be

TABLE I
THRESHOLDS FOR BENZENE AND METHYL SALICYLATE (IN MOLAR RATIOS)

O	Benzene		methyl salicylate
	0.2%	0.4%	
AB	1.7×10^{-4}	1.9×10^{-4}	4.1×10^{-7}
SH	1.6×10^{-4}	1.8×10^{-4}	5.4×10^{-6}
FNJ	1.8×10^{-4}	1.5×10^{-4}	3.6×10^{-6}
MHJ	1.1×10^{-4}	1.3×10^{-4}	3.6×10^{-8}
Mean	1.5×10^{-4}	1.6×10^{-4}	2.6×10^{-6}

of interest. First of all, one must know the vapor pressure of the odorous substance. The best source is probably Timmermans.⁵ If only the boiling point is known, it is possible to estimate the vapor pressure (as indeed was done for methyl salicylate in calculating the thresholds in this note).⁶ Now, from the densities and the molecular weights of the odor and the mineral oil, one can find the molar ratio of the solution, and hence calculate the vapor pressure of the substance as diluted (the vapor pressure of the mineral oil, which is very low, is neglected), by multiplying the vapor pressure by the mole fraction. The molar ratio of odor in the absorption bottle is then determined by reference to the air-pressure, and finally, the molar concentration of the threshold-blast is found from the volume of odorous air in the total mixture from which the blast is delivered.

⁴ Obtained respectively from the General Chemical Company and the Eastman Kodak Company.

⁵ Jean Timmermans, *Physico-Chemical Constants of Pure Organic Compounds*, 1950.

⁶ G. W. Thomson, Determination of Vapor Pressure, Chapt. V in Arnold Weissberger (Ed.), *Technique of Organic Chemistry*, Vol. I, Part 1, *Physical Methods of Organic Chemistry*, 1949, 228 (Formula 97).

Since the actual number of odorous molecules reaching the olfactory epithelium must depend upon a variety of factors, few of them the same for any two methods, no guess at the number of molecules involved in a barely supra-threshold stimulus will be attempted here, nor do comparisons with thresholds obtained by other methods have much meaning. If an apparatus can provide stable and homogeneous conditions with which to work, and if no variables other than stimulus-concentration, stimulus-substance, and *O* determine the thresholds, it has fulfilled its function. It appears that the apparatus herein described meets these criteria.

APPARATUS NOTES

THE BARNES-CZERNY APPARATUS FOR DEMONSTRATING QUANTAL FLUCTUATIONS IN VISION

The purpose of this note is to call attention to the Barnes-Czerny apparatus which enables the student to see quantal fluctuations as a scintillating dot-field similar to the twinkling of stars at night when atmospheric turbulence is high.¹

The apparatus itself is simple. It consists of an aluminum rectangular sheet 18 x 23 in., perforated with a set of holes, $\frac{1}{4}$ in. in diameter, spaced at the intersections of a 1-in. grid. The perforated sheet is painted flat black and mounted in an opening in a wall flush against a sheet of flashed opal glass. A 6-w., 120-v., candelabra base, bulb (GE6S6) is placed 4 ft. back of the opal glass in an adjoining lightproof room. It is not important whether the room is uniformly black or white. The bulb is connected to a variable auto-transformer (*Variac* 0-130 v., 5 amp. max.) which is in the room where S can look at the apertures in the aluminum sheet. He is seated about $2\frac{1}{2}$ ft. in front of the perforated plate with the transformer within convenient reach. He is told that the room will be suddenly darkened. He is asked to adjust the Variac until he can just see and just fail to see the dot-field. He is asked to describe what he sees. He discovers or can be primed to discover the course of dark-adaptation through the necessity of adjustment of the Variac control towards ever-lowering voltages for awhile. Then he finds that the setting of the control remains within a region—his terminal level of dark adaptation. The almost incredible sensitivity of the eye becomes apparent to the student when, after the experiment, he finds that the bulb was operated on from 4 to 12 v. with its filament barely glowing. He may also discover but he will more often have to be told to notice that the set of dots he looks at directly are not visible—the fact of foveal night blindness. Finally, however, most people report the peculiar appearance of the dot-field—it *twinkles*. Some Ss may have to move closer or farther from the apertures to see the twinkling best.

It is this observation of twinkling that Barnes and Czerny used to determine the average number of quanta required to excite night vision.

Other theoretical explanations can be offered—Crozier's neural fluctuations, scanning mechanisms, etc.—depending upon the level of the students.

There is another observation. The entire dot-field shows overall a slow waxing and waning in brightness. It is possible that this last observation may be correlated with the Traube-Hering waves.²

To clarify the quantum concept the author has found the following macroscopic analogy useful. The light quanta coming through the perforated plate can be likened to grains of sand falling on the top of a set of test tubes arranged in the

¹ R. B. Barnes and M. Czerny, *Lässt sich ein Schroteffekt der Photonen mit dem Auge beobachten?*, *Zsch. f. Physik.*, 79, 1932, 436-449.

² R. S. Woodworth, *Experimental Psychology*, 1938, 701-702.

same pattern as the perforated plate. The openings in the test tubes correspond to the apertures in the plate. The light intensity—quanta per second per unit area—corresponds to the rate of fall and random density of the sand grains. (Of course, the rate variable is not appropriate for light whose velocity is constant for a given medium—this is one instance where the analogy is not complete.) If just a few grains of sand are falling the number of grains collected in each test tube with short duration samples will vary from tube to tube more than when more grains of sand are falling. It is this variability from tube to tube that corresponds to the variability in the number of quanta passing through the apertures of the perforated plate. It is this variability which the dark adapted eye sees as a twinkling of the dot-field.

Yale University

LLOYD HENRY BECK

A GRID FOR ADMINISTERING SHOCK

The use of a conventional grid for administering shock to unrestrained animals (rats) has the disadvantage that an S may stand on two bars of the grid which are on the same side of the electrical circuit, and thus avoid the shock.¹ This possibility may be eliminated by use of the circuit diagramed in Fig. 1. The device makes use of a motor driven commutator which connects one bar of the grid at a time to the

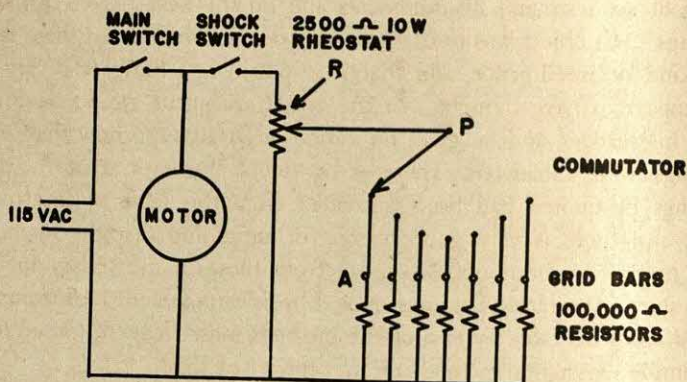


FIG. 1. CIRCUIT FOR GRID

This circuit provides for shock between any two bars of the grid.

center pole of a rheostat; repeating the cycle several times a second. All bars are connected to the other side of the circuit through individual 100,000 ohm resistors. An S making contact with any two or more bars will receive a shock every time any one of these bars is connected. The rheostat may be adjusted to deliver a shock ranging from 0 to about 0.5 m. amp. assuming a skin resistance of 100,000 to 200,000 ohms. The upper limit, 0.5 m. amp., is a strong shock for a rat.

Experimental results indicate that rats will learn an avoidance response appreciably faster when the present device, rather than a conventional grid, is used.

University of Wisconsin

L. BENJAMIN WYCKOFF

HORACE A. PAGE

¹ K. F. Muenzinger and F. C. Walz, An examination of electrical current stabilizing devices for psychological experiments. *J. Gen. Psychol.*, 10, 1934, 477-482; B. F. Skinner and S. L. Campbell, An automatic shocking grid apparatus for continuous use, *J. Comp. & Physiol. Psychol.*, 40, 1947, 305-307.

NOTES AND DISCUSSIONS

THE EARLIEST FORMULAE USED IN FACTOR ANALYSIS

Spearman in his famous paper of 1904 quotes a table of the inter-correlation of six tests and measurements and on the next page a list of 'g' loadings.¹ His object was to expound the then novel idea that there is only one kind of intelligence, and that it is possible to measure it. He does not appear to have thought that the actual arithmetic would have interested his readers, and he gives no indication at all as to how these factor loadings were calculated. There is no doubt, however, that the factor loadings he quoted had been calculated from this table of intercorrelations, and there is very good reason for supposing that the method he used must have been not dissimilar from those in use today. In Table I are shown the factor loadings quoted by Spearman and the loadings obtained from his table by two of the methods most frequently used today; the simple summation method and the centroid method.

TABLE I
SPEARMAN'S FACTOR LOADINGS WITH THOSE OBTAINED FROM HIS TABLE
BY MODERN METHODS

Loadings	Classics	French	English	Math.	Pitch	Music
Spearman's (1904)	.99	.92	.90	.86	.72	.70
Obtained by simple summation	.96	.89	.80	.75	.67	.65
Obtained by the centroid method	.91	.88	.82	.77	.70	.68

Spearman quotes two separate 'g' loadings for discrimination of pitch. Both could not have been calculated from his table of inter-correlations; one at least must have been obtained from another source, as were four other loadings in the list. The 'g' loadings for discrimination of pitch are for 'cultured' and for 'uncultured' persons, and since school boys of 11-12 yr. of age could hardly be described as cultured, at least not in the musical sense, it is reasonable, to suppose that the loadings for 'uncultured' persons was the one he obtained from the correlations of the scores of his schoolboys.

¹ Charles Spearman, "General intelligence" objectively determined and measured, this JOURNAL, 15, 1904, 202-292, esp. 275 f.

When there are only six tests and the matrix is not a very close approximation to rank one, different methods of factorizing are likely to give somewhat different loadings. Spearman's method may be slightly less accurate than the other two, but the difference is not very great, and when it is remembered that this is the first factor analysis ever made, the results must seem surprisingly good.

There is no evidence that Spearman ever at any time used a method of factorizing which employed guessed communalities, whereas there is reason to believe that he was opposed to this method. In a single-factor factorization where the communalities are the squares of the factor loadings, to use guessed communalities would be to run the risk of being accused of assuming what it was required to prove.

The writer has worked on Spearman's correlations using different factorizing methods that do not use guessed communalities, in the hope of finding one that gives the same loadings as those quoted by Spearman. Some results are shown in Table II below.

Formula 2 is the earliest known Spearman formula, which will be mentioned later, and Formula 3 is an obvious simplification of it. The other formulae are expressed in the same notation, where

a_s = the sum of column s , the communalities being omitted,

2Σ = the sum of all the correlations in the table, the communalities being omitted, and

n = the number of tests.

Formula 4 is a very unlikely one, but it does give values that are close to Spearman's.

It is obvious that Spearman cannot have used any of these formulae; none of the loadings obtained agrees sufficiently closely with his values. Since he has quoted his correlations rounded off to two figures, an exact agreement is not to be expected, but that should not cause a disagreement of more than about 0.01 in any of the loadings.

In an endeavor to find a clue to the method Spearman had used, a careful study was made of the relevant passages in his 1904 paper and of the references to this work in his paper of 1906.² There seems to be good reason to suppose that Spearman had attached greater importance to the correlations between Classics, French, English, and Mathematics than to the others. In the discussion of hierarchical order on pages 274 and 275 of his 1904 paper he discusses these measurements first, and

² Spearman, Die Korrelation zwischen verschiedenen geistigen Leistungsfähigkeiten, *Zsch. f. Psychol.*, 44, 1906, 87.

only later mentions the measurements of discrimination of pitch and musical ability, which actually fit into the hierarchy quite as well as the others. In the table of correlations on page 275 he quotes the reliabilities of the scores for Classics, French, English, and Mathematics but not for the other two measurements. Lower on the same page he says that the

TABLE II
CORRELATIONS WITH DIFFERENT METHODS OF FACTORIZING

Formula	Classics	French	English	Math.	Pitch	Music
(1) Spearman 1904	.99	.92	.90	.86	.72	.70
(2) $\frac{a_s}{\sqrt{2\Sigma - a_s}} \cdot \sqrt{\frac{n-2}{n-1-\frac{na_s}{2\Sigma}}}$	.95	.88	.81	.75	.67	.65
(3) $\frac{a_s}{\sqrt{2\Sigma - a_s}}$	.94	.87	.80	.75	.68	.66
(4) $\frac{a_s}{\sqrt{2\Sigma - a_1}}$	.97	.91	.84	.80	.72	.70
(5) $\frac{a_s}{\sqrt{2\Sigma}}$	.84	.79	.73	.69	.63	.61
(6) $\frac{a_s}{\sqrt{2\Sigma}} \cdot \sqrt{\frac{n}{n-1}}$	.92	.87	.80	.76	.69	.67
(7) $\frac{a_s}{\sqrt{2\Sigma}} \cdot \sqrt{\frac{n+1}{n}}$	.91	.85	.79	.75	.68	.66
(8) $\frac{a_s}{\sqrt{2\Sigma}} \cdot \frac{n}{n-1}$	1.01	.95	.88	.83	.75	.73
(9) $\frac{a_s}{\sqrt{2\Sigma}} \cdot \frac{n+1}{n}$	.98	.92	.85	.81	.73	.71

probable errors vary from about 0.01 for the 'ordinary studies' to 0.03 for music. In his 1906 paper he again discusses Classics, French, English, and Mathematics together, and mentions music and discrimination of pitch almost as an after-thought. This paper states that there were 36 pupils in the group, but the appendix to the 1904 paper (pages 290 and 291) makes it clear that only 22 of them had taken music as a school subject. It appears from a statement in the 1906 paper that there were no marks for music; the pupils were arranged in order of ability by the music master.

Whatever may have been Spearman's reasons for attaching greater importance to Classics, French, English, and Mathematics, there is no doubt that he did. This fact suggested to the writer that, if Spearman regarded these measurements as being more reliable, he may have thought

TABLE III
MATRIX REDUCED TO INTERCORRELATION OF FOUR MEASUREMENTS

Formula	Classics	French	English	Math.	Sum of deviations from Spearman's values
(1) Spearman 1904	.99	.92	.90	.86	
(2) Mean of $\sqrt{\frac{r_{as} \cdot r_{bs}}{r_{ab}}}$	.95	.87	.83	.75	.25
(3) $\frac{a_s}{\sqrt{2\Sigma - a_s}} \cdot \sqrt{\frac{n-2}{n-1 - \frac{na_s}{2\Sigma}}}$	.96	.85	.80	.75	.29
(4) $\frac{a_s}{\sqrt{2\Sigma - a_s}}$	.92	.85	.82	.78	.30
(5) $\frac{a_s}{\sqrt{2\Sigma}}$	.79	.74	.71	.69	.74
(6) $\frac{a_s}{\sqrt{2\Sigma}} \cdot \sqrt{\frac{n}{n-1}}$	.91	.85	.82	.79	.30
(7) $\frac{a_s}{\sqrt{2\Sigma}} \cdot \sqrt{\frac{n+1}{n}}$	.88	.83	.80	.77	.39
(8) $\frac{a_s}{\sqrt{2\Sigma}} \cdot \frac{n}{n-1}$	1.05	.99	.95	.91	.23
(9) $\frac{a_s}{\sqrt{2\Sigma}} \cdot \frac{n+1}{n}$	.99	.92	.89	.86	.01

that a more accurate estimation of their 'g' loadings could be obtained from them alone; the pitch and music correlations being omitted from the table.

The matrix reduced to the intercorrelations of these four measurements was factorized by the more likely methods, and the results are shown in Table III.

It will be seen that Formula 9 gives loadings that are practically identical with those quoted by Spearman. Three agree exactly and one differs

by only 0.01. Since these loadings have been calculated from correlations that have been rounded off to two figures, no closer agreement could be expected.

Factorizing formulae which give practically identical loadings with a large table of correlations will give quite considerable differences in the loadings with a table of the correlations of only four tests. With so small a table, a formula would be expected to give considerable deviations from Spearman's values *unless it was the formula that Spearman used*. This is exactly the picture presented by Table III. The sum of the deviations from Spearman's values are shown in the last column. In the case of Formula 9 it is 0.01, the next nearest is 0.23.

It may be argued that this proves nothing, since if a large enough number of formulae are tried, sooner or later one will be found to give values close to Spearman's. Against this is the fact, shown by Table III, that 'wrong' formulae do not give values that are close to Spearman's, they give values that are very different. Also, it is not possible to go on trying different formulae; the number of possible formulae is very limited.

There are two other pieces of evidence in favor of Formula 9. If Spearman used this formula to obtain loadings for Classics, French, English, and Mathematics, it is equally likely that he used it to obtain the loadings for Pitch and Music, by factorizing all six of the tests. If he did, he would have been expected to have quoted the loadings found with this formula and shown in the last row of Table II. These are 0.73 and 0.71, and the loadings he quotes are 0.72 and 0.70. There is a discrepancy of 0.01 in each case. This is too close to be a coincidence.

The other piece of evidence is that the earliest known Spearman formula (Formula 2 of Table II, and Equation [i] following) is of the form of an approximate formula with a correction factor, and Formula 9 of Table II is also of this form.

All this is very far short of proof, but at least it seems extremely probable that in the first single-factor factorization ever to be made, the formula used was

$$m_s = \frac{a_s}{\sqrt{2\Sigma}} \cdot \frac{n+1}{n}.$$

In the early days Spearman did not publish any of his factorization formulae, but in one of his letters written in 1909, that is now among the Spearman collection at the (British) National Institute of Industrial Psychology, he says that there are various formulae that can be used, depending on the degree of exactness aimed at, but that the results are

practically the same in any case. In the same letter he gives one of his formulae; (it is Formula 2 of Table II). He stated it as follows:—

$$r_{st} = m_s m_t$$

$$m_s = \frac{a_s}{\sqrt{2\Sigma - a_s}} \times \sqrt{\frac{n-2}{n-1 - \frac{na_s}{2\Sigma}}} \dots\dots\dots [1]$$

a_s = the sum of all the correlations with the performance 's'

2Σ = the sum of all the different correlations altogether

n = the number of performances

m_t has a value analogous to m_s

A little later Burt published an account of some experimental work in which he quoted this formula, attributing it to Spearman.³ It is possible that more of Spearman's correspondence or personal memoranda may come to light and some others of his early formulae may be discovered, but until they do, this is the first factorization formula of which we have any record.

It will be obvious that the quantity

$$\sqrt{\frac{n-2}{n-1 - \frac{na_s}{2\Sigma}}}$$

tends to unity when n is large, and is of the nature of a correction factor. The remainder of the formula can be used without it, without any great loss of accuracy. The formula is, in fact, surprisingly accurate. As an illustration an artificial matrix of the correlations of six "tests" having factor loadings of 0.8, 0.7, 0.6, 0.4 and 0.3, when factorized with this formula, gives the loadings shown in Table IV. For comparison, the loadings obtained by the centroid formula are also shown.

TABLE IV
COMPARISON OF THE LOADINGS

True loadings	.8	.7	.6	.5	.4	.3
Spearman's formula with correction factor	.797	.704	.607	.501	.406	.303
Spearman's formula without correction factor	.761	.648	.600	.511	.416	.318
Centroid method	.758	.703	.620	.535	.437	.337

³ Cyril Burt, Experimental tests of general intelligence, *Brit. J. Psychol.*, 3, 1909, 162.

The derivation of the first part of the formula is fairly obvious. If the matrix of correlations is written in the form of the factor products,

$$\begin{array}{ccccccc} & & m_1m_2 & m_1m_3 & m_1m_4 & \cdots & \\ m_2m_1 & - & & m_2m_3 & m_2m_4 & \cdots & \\ m_3m_1 & m_3m_2 & - & & m_3m_4 & \cdots & \\ m_4m_1 & m_4m_2 & m_4m_3 & - & & \cdots & \\ \text{etc.} & \cdots & & & & & \end{array}$$

then in Spearman's notation

$$\begin{aligned} a_1 &= m_1m_2 + m_1m_3 + m_1m_4 \cdots \\ &= m_1(m_2 + m_3 + m_4 \cdots) \cdots \cdots \cdots \quad [\text{ii}] \end{aligned}$$

and similarly

$$\begin{aligned} a_2 &= m_2(m_1 + m_3 + m_4 \cdots) = m_2(m_2 + m_3 + m_4 \cdots) + m_2m_1 - m_2^2 \\ a_3 &= m_3(m_1 + m_2 + m_4 \cdots) = m_3(m_2 + m_3 + m_4 \cdots) + m_3m_1 - m_3^2 \\ a_4 &= m_4(m_1 + m_2 + m_3 \cdots) = m_4(m_2 + m_3 + m_4 \cdots) + m_4m_1 - m_4^2 \end{aligned} \left. \cdots \right\} \quad [\text{iii}]$$

etc. . . .

The sum of all the terms on the right in heavy type is likely to be a very small quantity, so that, summing all the Equations [iii]

$$a_1 + a_2 + a_3 + \cdots = (m_2 + m_3 + m_4 + \cdots)^2 + \text{a small quantity}$$

Thus, the first part of Spearman's formula

$$\begin{aligned} \frac{a_1}{\sqrt{2\Sigma - a_1}} &= \frac{a_1}{\sqrt{a_2 + a_3 + a_4 + \cdots}} \\ &= \frac{m_1(m_2 + m_3 + m_4 \cdots)}{\sqrt{(m_1 + m_2 + m_3 \cdots)^2 + \text{a small quantity}}} \\ &= m_1 (\text{approx.}) \end{aligned}$$

It is quite a close approximation as can be seen from Table IV.

The derivation of the first part of the formula is obvious, but how Spearman obtained the rest of the formula is not so certain. The small quantity that was neglected contains the terms m_2^2 , m_3^2 , m_4^2 , etc., which cannot be calculated. The question is how Spearman obtained an approximate value for them. The writer suggests the following derivation as the most likely, as it employs two devices that Spearman has used in other mathematical work.

The first is a method of summing all the correlations of a matrix that do not involve some particular test 's.' It is

$$(\text{Sum of all correlations not involving 's'}) = 2\Sigma - 2a_s$$

where Σ and a_s have the same meaning as in Equation [i]. Spearman used this device in the factorization formula usually associated with his name.

The other device is that of exchanging the arithmetic and geometric means of quantities of similar magnitude. This device can be used to obtain approximate values for the unknown quantities m_2^2 , m_3^2 , m_4^2 , etc., in terms of known quantities. For instance,

$$\begin{aligned}\frac{1}{2}(m_2^2 + m_3^2) &= \sqrt{m_2^2 m_3^2} \text{ (approx.)} \\ &= r_{23} \text{ (approx.)}\end{aligned}$$

In the case of three such quantities

$$\begin{aligned}m_2^2 + m_3^2 + m_4^2 &= \frac{1}{2}[(m_2^2 + m_3^2) + (m_3^2 + m_4^2) + (m_4^2 + m_2^2)] \\ &= \frac{1}{2}[2r_{23} + 2r_{34} + 2r_{42}].\end{aligned}$$

If this device is used to average all the possible values of $(m_2^2 + m_3^2 + m_4^2 \dots)$ throughout the whole matrix, we get

$$\begin{aligned}(m_2^2 + m_3^2 + m_4^2 \dots) &= \frac{1}{n-2} \text{ (sum of all the correlation coefficients in the matrix} \\ &\quad \text{that do not involve test 1)} \\ &= \frac{1}{n-2} (2\Sigma - 2a_1) \dots \dots \dots \text{ [iv]}\end{aligned}$$

Summing Equations [iii] shows that

$$2\Sigma - a_1 = (m_2 + m_3 + m_4 \dots)^2 + m_1(m_2 + m_3 + m_4 \dots) - (m_2^2 + m_3^2 + m_4^2 \dots).$$

Substituting Equations [ii] and [iv]

$$2\Sigma - a_1 = (m_2 + m_3 + m_4 \dots)^2 + a_1 - \frac{1}{n-2} (2\Sigma - 2a_1).$$

So that,

$$(m_2 + m_3 + m_4 \dots)^2 = (2\Sigma - a_1) - \left[a_1 - \frac{1}{n-2} (2\Sigma - 2a_1) \right] \dots \dots \dots \text{ [v]}$$

If, as the writer believes, Spearman did use these two artifices and obtained the relationship expressed in Equation [v] the derivation of Spearman's factorization formula, Equation [i] above, is fairly obvious.

Beginning with the identity;

$$m_1 = \frac{m_1(m_2 + m_3 + m_4 \dots)}{\sqrt{(m_2 + m_3 + m_4 \dots)^2}}$$

and substituting Equation [ii] in the numerator and Equation [v] in the denominator.

$$\begin{aligned} m_1 &= \frac{a_1}{\sqrt{(2\Sigma - a_1) - \left[a_1 - \frac{1}{n-2} (2\Sigma - 2a_1) \right]}} \\ &= \frac{a_1}{\sqrt{(2\Sigma - a_1) \left[1 - \frac{a_1 - \frac{1}{n-2} (2\Sigma - 2a_1)}{2\Sigma - a_1} \right]}} \\ &= \frac{a_1}{\sqrt{(2\Sigma - a_1) \left[1 - \frac{-2\Sigma + na_1}{(n-2)(2\Sigma - a_1)} \right]}} \\ &= \frac{a_1}{\sqrt{(2\Sigma - a_1) \left[1 - \frac{-1 + \frac{na_1}{2\Sigma}}{n-2} \right]}} \quad (\text{approx.}) \\ &= \frac{a_1}{\sqrt{(2\Sigma - a_1) \left[\frac{n-1 - \frac{na_1}{2\Sigma}}{n-2} \right]}} \end{aligned}$$

or in the form in which Spearman presented it

$$m_1 = \frac{a_1}{\sqrt{2\Sigma - a_1}} \cdot \sqrt{\frac{n-2}{n-1 - \frac{na_1}{2\Sigma}}}$$

It is an approximation and not an exact value, but just how good an approximation it is can be seen by comparing the first and second rows of the Table IV.

National Institute of Industrial Psychology
London, England

D. F. VINCENT

METHODS FOR COMPUTING UNCERTAINTIES

This note should be regarded as a footnote to Newman's valuable paper on the computation of uncertainties and autocorrelation coefficients.¹ It will propose a simplification in tabulation techniques for computing uncertainties in binary data, and will discuss how to compute uncertainties in certain kinds of non-binary data.

Simplified tabulation of binary data. The basic requirements for computation of uncertainties is a table which shows the frequencies of occurrence of various sequences of events. Often these events are binary, *i.e.* at any moment only two alternative events can occur. For instance, a rat in a T-maze in the absence of reinforcement may turn right or left; a rat in a Skinner box may press the bar or approach the food-tray; a human being may judge a given tone as louder or softer than a standard. Let us call the two possible events in such a situation A and B . It often happens that the choice between these two events is influenced by the sequence of such events which has already occurred. If the longest sequence length we wish to consider is four, then a table must be prepared from the data showing $N_{AAAA}, N_{AABA}, \dots, N_{BBBB}$, where N_{AAAA} is the number of times the sequence of events $AAAA$ occurred.

Once the number of occurrences of all possible sequences of length s is counted from the data, the number of occurrences of sequences $s - 1$, $s - 2, \dots, 1$ long can be determined by addition. There are, however, two ways to do this, as Table I illustrates. Table IA shows some of Newman's data tabulated as he tabulated it; Table IB is the same data tabulated in the other possible way. In Table IA, N_{ABA} is found by adding N_{ABAA} and N_{ABAB} ; in Table IB, N_{ABA} is found by adding N_{AABA} and N_{BABA} . Notice that the value of N_{ABA} is the same in each case, as it should be. In short, it does not matter whether N_{ABA} is determined by summing the number of occurrences of the two sequences in which ABA are the first three symbols, or the number of occurrences of the two sequences in which ABA are the last three symbols.

* This research was done in cooperation with the Systems Coordination Division, Naval Research Laboratory under Contract N5-ori-166, Task Order I, between the Office of Naval Research and The Johns Hopkins University. This is Report No. 166-I-169, Project Designation No. NR 507-470, under that contract.

¹ E. B. Newman, Computational methods useful in analyzing series of binary data, this JOURNAL, 64, 1951, 252-262. For more complete presentations of the fundamental concepts underlying computations of this sort, see C. E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.*, 27, 1948, 379-423, 623-656; and F. C. Frick and G. A. Miller, Statistical behavioristics and sequences of responses, *Psychol. Rev.*, 56, 1949, 311-324.

This fact makes possible a considerable reduction in tabulating labor. A system of equations which looks like this can be set up:

$$N_{ABAA} + N_{ABAB} = N_{AABA} + N_{BABA} (= N_{ABA}) \dots \dots \dots [1]$$

For binary data, there will be 2^S different sequences of length S which must be tabulated. Since there are $2^S/2$ equations like Equation [1], it seems as if it should be possible to count from the data only half of the 2^S

TABLE IA

NEWMAN'S DATA ON VOWEL AND CONSONANT FREQUENCIES AS HE TABULATED IT

AAAA	21	AAA	90		
AAAB	69			AA	274
AABA	155	AAB	184		
AABB	29				A 601
ABAA	143	ABA	263		
ABAB	120			AB	327
ABBA	56	ABB	64		
ABBB	8				
BAAA	69	BAA	184		
BAAB	115			BA	327
BABA	108	BAB	143		
BABB	35				B 399
BBAA	41	BBA	64		
BBAB	23			BB	72
BBBA	8	BBB	8		
BBBB	0				

sequences, and determine the frequency of occurrence of the rest by solving this set of simultaneous equations. Actually, however, this is not possible because the equations associated with, for instance, N_{AAA} and N_{BBB} take the form:

$$N_{AAAA} + N_{AAAB} = N_{A AAA} + N_{B AAA} (= N_{AAA}) \dots \dots \dots [2]$$

which reduces to

$$N_{AAAB} = N_{BAAA} \dots \dots \dots [3]$$

and so gives no information about N_{AAAA} at all. Consequently, it turns out to be necessary to supply $(2^s/2) + 1$ counts of sequences before the simultaneous equations can be solved. The extra sequence can be counted directly from the data, and this is in general the best way to supply it. But it is also possible to add one more equation to the set of simultaneous equations:

$$N_{AAAA} + N_{AAAB} + N_{AABA} + \dots + N_{BBBB} = N_{tot}, \dots [4]$$

where N_{tot} is the total number of symbols in the entire sequence minus $(s - 1)$. Use of this equation eliminates all internal checks on the correct-

TABLE IB
ALTERNATIVE TABULATION OF NEWMAN'S DATA

AAAA	21	AAA	90	AA	274	
BAAA	69					
ABAA	143	BAA	184			
BBAA	41					A 601
AABA	155	ABA	263			
BABA	108			BA	327	
ABBA	56	BBA	64			
BBBA	8					
AAAB	69	AAB	184			
BAAB	115			AB	327	
ABAB	120	BAB	143			
BBAB	23					B 399
AABB	29	ABB	64			
BABB	35					
ABBB	8	BBB	8	BB	72	
BBBB	0					

ness of tabulation, and so is dangerous unless there is good reason to be sure that all counts made from the data are correct. When tabulation is being done by machine, however, it may be much more convenient to make $2^s/2$ counts than $(2^s/2) + 1$, since sometimes the extra count may require as much time to make as all $2^s/2$ (as for example in a machine with 16 counters when $s = 5$, and so there are 32 sequences to be tabulated). In

such cases using Equation [4] to provide the extra count may be justified.

More complex equations based on the various different ways of adding the sequences of length s to get sequences of length $s - 2$, $s - 3$, etc.,

TABLE II
IMAGINARY T-MAZE DATA (ONLY HALF THE TABLE IS GIVEN)

(5)		(4)	(3)	(2)	(1)
L+L+L	100	L+L+	150		
L+L+R	50		L+L	300	
L+L-L	145	L+L-	150		
L+L-R	5			L+	400
L+R+L	30	L+R+	50		
L+R+R	20		L+R	100	
L+R-L	30	L+R-	50		
L+R-R	20				L 800
L-L+L	150	L-L+	190		
L-L+R	40		L-L	380	
L-L-L	185	L-L-	190		
L-L-R	5			L-	400
L-R+L	8	L-R+	10		
L-R+R	2		L-R	20	
L-R-L	8	L-R-	10		
L-R-R	2				

Uncertainties calculated by Newman's method: $U_1=0.7220$, $U_2=1.0000$, $U_3=0.6333$, $U_4=0.9999$, $U_5=0.5892$

Uncertainties calculated from Equation [6]: $U_1=0.7220$, $U_3=0.6333$, $U_5=0.5888$

may be written, but they do not reduce the number of sequences which must be counted from the data any further.

Difficulties at beginning and end of sequence. Any tabulation of sequences of length s must result in a sum over all possible sequences which is $s - 1$ smaller than the total number of symbols in the sequence. Thus, the sequence *AABAAB* contains only 5 different sequences of length 2; 4 of length 3, etc. That is why N_{tot} was defined above as being the total number of symbols minus $(s - 1)$.

This fact can lead to minor discrepancies between the two ways of

summing shown in Table I. For instance, the sequence *AABAAB* contains 2 *AAB*s, 1 *ABA*, and 1 *BAA*. If we try to find N_{AA} by adding N_{AAA} and N_{AAB} , we get 2, which is the right answer; but if we try to find N_{AA} by adding N_{AAA} and N_{BAA} , we get only 1. On the other hand, if we try to find N_{AB} by adding N_{ABA} and N_{ABB} , we get one short of the right answer, while adding N_{AAB} and N_{BAB} gives the right answer.

In general, when computing frequencies of occurrence of sequences by addition, r counts which could have been made from the data will be omitted if sequences of length $s - r$ are being computed. If the method of addition in Table IA is used, these r counts will be the first $s - 1$ symbols in the total sequence; if the method of addition in Table IB is used, they will be the last $s - 1$ symbols. If the frequencies from more than one subject or more than one set of choices from the same subject are summed, then Pr elements are lost, where P is the number of different sets of data being included in the summation. This loss is unimportant so long as N_{tot} is very large relative to s , as it always is in situations where this kind of analysis is reasonable.

The small discrepancies which this fact can produce between the two otherwise equivalent methods of summing frequencies of occurrences of longer sequences to find those of shorter ones are its only practical effect. No such discrepancies occurred in Newman's data, reproduced in Table IA and IB. This could be either because the same three symbols occurred at the beginning and at the end of his data, or because he used the device of repeating the first four symbols at the end of the tape into which his data were punched. This avoids the discrepancies in summations, but distorts slightly the counts of the longest sequences.

Computing uncertainties from non-binary data. Newman says, "The particular procedures developed here have one important limitation: they are applicable only to sequences of *binary* data."² This statement, although perfectly correct when applied to the machines Newman discusses and to the autocorrelation computation, is not correct when applied to his method of computing uncertainties. The essence of his method is a subtractive process based on the equation

$$U_s = - \sum p(i,j) \cdot \log_2 p(i,j) - (U_{s-1} + U_{s-2} + \dots + U_1) \dots [5]$$

This equation, and therefore the subtractive procedure based on it, are perfectly general, and not restricted to the binary case.³

² Newman, *op. cit.*, 252.

³ See p. 169.

Bi-binary data. A special problem arises with certain kinds of data which look as if they were binary, but really might be called bi-binary. Consider a T-maze experiment. The rat may turn right (R) or left (L), and he may be rewarded (+), or reward may be withheld (-). A set of ten trials might be represented as follows: L + L - R + L - L + R - L + L + L - L. Such data could be arranged as in Table II. In this form, it looks binary. (Note, however, that equations like Equation [1] do not apply.) Can uncertainties be calculated by Newman's method for this kind of data?

Actually there are two different kinds of binary events occurring here, L or R, and + or -. Each of these events can be predicted from prior sequences of events, and the average uncertainty of each of these predictions can be calculated. In Table II, U_1 , U_3 , and U_5 are uncertainties of predicting responses based on knowledge of 0, 1, and 2 previous responses (and their consequent rewards or non-rewards), while U_2 and U_4 are uncertainties of predicting rewards or non-rewards. All of these uncertainties are correct, and have been calculated by Newman's method. Note that U_2 is larger than U_1 and U_4 is larger than U_3 . This is an apparent violation of the rule that uncertainties cannot be increased by more information. The violation is not real, however, since the event being predicted in U_2 and U_4 is different from that being predicted in U_1 and U_3 .

In T-maze and other similar experiments, we are not usually interested in the uncertainties of predicting rewards. Consequently, use of Newman's subtractive procedure, although it results in correct answers, requires the calculation of unnecessary uncertainties. In such cases, it is often simpler to

^a Very recently, W. J. McGill has published a new and superior method of calculating uncertainties, based on the equation:

$$Q_{ijk} \dots = 1/N_{tot} \sum_{ijk} \dots N_{ijk} \dots \log_2 N_{ijk} \dots [7]$$

where N_{ijk} is the number of occurrences of sequence ijk . To use this method on the data in Table IA, for instance, simply calculate $N \log_2 N$ for each raw number in the table. The sum of the $N \log_2 N$ values associated with the first column will be Q_{ijk} , the sum associated with the second column will be Q_{jk} , and the sum associated with the third column will be Q_k . Then,

$$U_1 = Q - Q_k, \text{ where } Q = \log_2 N_{tot}; U_2 = Q_k - Q_{jk}; U_3 = Q_{jk} - Q_{ijk}; \text{ etc.}$$

This method, which is applicable to any kind of data, binary or otherwise, is much faster than Newman's method. (It is another way of applying the same subtractive procedure.) Its only inconvenience is that it requires a table of $N \log_2 N$ values, and no such table is conveniently available as yet. It is an outgrowth of a fascinating new method for dealing with transmitted information in a manner analogous to analysis of variance. See W. J. McGill, Multivariate transmission of information and its relation to analysis of variance, HFORL Report No. 32, Human Factors Operations Research Laboratories, Air Research and Development Command, USAF.

calculate uncertainties directly from the fundamental formula:

$$U_s = -\sum_{i,j} P_i p_i(j) \log_2 p_i(j) \dots \dots \dots [6]$$

where P_i is the probability of being in state i , and $p_i(j)$ is the probability, given that state i exists, of event j following it. Since this formula gives U_s directly, the extra uncertainties need not be calculated. In order to use this formula, the method of addition shown in Table IA rather than that in Table IB must be followed. For all non-binary data it is essential to follow the rule that when summing, each shorter sequence must represent the antecedent of the last symbols in the longer sequences which were summed to find it. This holds true regardless of the method of calculation used. It is not essential for binary data only so long as Newman's subtractive procedure is used.

If data like these are to be tabulated by machine, they must not, of course, be fed into the machine in the form they take in Table II. To most such tabulating machines, each event must be of the same sort as the event before it; such machines are not equipped to distinguish between responses and reinforcements. This problem can easily be dealt with by coding the data in quaternary form. Thus, L+ might be A; L-, B; R+, C; and R-, D. Then the data can be counted by a machine equipped to count quaternary data, and subsequently can be reconverted to the bi-binary form of Table II. Equations for such quaternary data similar to Equation [1] for binary data can be written, but do not save much tabulating time, since $3/4 \cdot 2^8 + 1$ counts must be supplied from the data before they can be solved. The same is true of bi-binary data. Such equations may be useful, however, as internal checks on the correctness of tabulation.

The Johns Hopkins University

WARD EDWARDS

RECURRENT IMAGES: NEW INSTANCES AND A SUMMARY OF THE OLDER ONES

After a family excursion into a blackberry patch, my wife reported to me as she closed her eyes after retiring that she could see beautiful blackberries, just perfect for picking, hanging in great profusion on bushes. Upon closing my eyes I discovered that I too saw them. The images greatly impressed me for they were neither after-images in the usual sense nor were they memory images. The images were positive and appeared to be located in the eyes rather than projected. They were

very vivid; they could be *seen* not just imagined as in the case of memory images. In this respect they were like the usual after-image. Introspectively they appeared to be retinal phenomena.

The berry-picking excursion occurred during a bright sunny day about eight hours before we retired. My wife and I both saw idealized images; the berries were large, purple-tinged, luscious, and profuse. In place of the almost prohibitive brier patch, the berries hung upon open shoots; they were easily accessible and all of them were ripe. During the excursion we had seen many small berries, green berries, and bird-pecked berries, but in our images not an imperfect berry appeared. The background was present but subdued in comparison to the berries. The green leaves of the bushes were seen but greatly reduced in number and saturation.

My wife reported that she had had similar images before under comparable circumstances. This was my first experience but I have since noted the phenomenon after other berry-picking excursions which include strawberries and blueberries as well as blackberries. Upon inquiry I found that other psychologists have had similar experiences and that they were as much at a loss to explain them as I was. This sent me to the literature to find a description of them for surely anything as real and as vivid as these images must have been experienced and described. My search was not very fruitful. I found Titchener's "recurrent images,"¹ and Warren's "delayed after-sensations,"² but their experiences, though similar, did not at first seem to be identical with mine. After communicating with Professor Dallenbach and receiving additional references from him, I am now convinced, however, that they are. My experiences are instances of recurrent images. Since many psychologists are not acquainted with this type of imagery, a review of the descriptions of it may be of value.

Titchener, who is the first to mention this type of imagery, wrote of it in 1915 as follows:

"All after-images are sensory in character. . . . So also are the *recurrent images*, those troublesome and haunting images to which most of us are subject at times: the tunes that run in our head and that we cannot get ride of, the rows of figures that obsess us after a long morning of calculation, the bright disk that keeps cropping up after we have spent several hours at the microscope. . . . *Not every one has recurrent images.*"³

Two years later, in 1917, Dallenbach reported the experience of re-

¹ E. B. Titchener, *A Beginner's Psychology*, 1915, 75.

² H. C. Warren, Some unusual visual after-effects, *Psychol. Rev.*, 28, 1921, 453-463.

³ Titchener, *op. cit.*, 75 (italics are boldface in the original).

current visual images of a chessboard and pieces which persisted for several days after playing his first game of blindfold chess.⁴

During this game he sat with his back to the board being used in play and, as an aid, gazed at a board, an *objet d'art*, hung on the wall in front of him. The "considerable mental strain" involved in keeping the game in mind as he gazed at the vertical board was apparently the source of the recurrent images of the board. Since this board had no pieces on it, the presence of them in the recurrent image is a bit baffling. Can imagination be the source of a recurrent image? The images of the pieces on the board were undifferentiated; they were all "like a checker." Since he had been an enthusiastic player of draughts since childhood, these imagined pieces were probably tied up with images of his early experiences. It is not clear, however, what proportion of the rich imagery reported in this study should be classified as recurrent imagery but certainly some of it must be.

In 1921, Warren reported some unusual visual after-effects that he had experienced. Although some of his experiences were definitely recurrent images, he made no reference to Titchener nor to Dallenbach.

During World War I, it was Warren's custom to trace, just before retiring, the day's progress in various parts of the battle front on a large wall map. The room was dark except for a movable electric light which intensely illuminated different parts of the map as he shifted the lamp. On several occasions he noticed after going to bed and closing his eyes that he experienced vivid, positive images "which bore considerable likeness to tracings on the map." Another experience, alike in kind, occurred after he had studied with a microscope three slides of the brain. After about 15 min. of sleep he awoke and upon again closing his eyes he saw a disk of light comparable to the field of the microscope and little black dots resembling the nerve cells. His images of the printed page and of cards after seeing them under bright illumination and retiring are also cases in point. None of these images (of the map, microscopic slides, printed page, or cards) were stationary; they changed and moved about very much as the sensations did in perception. Regarding these phenomena, Warren wrote, "I wish to emphasize strongly the *definiteness* of the figures seen, their *motion*, and the fact that they appeared unmistakably of *retinal* origin."⁵

In 1924, Dallenbach reported two instances of recurrent images which he experienced as the result of a 400-mile automobile trip.⁶ One a new type, was kinesthetic, and the other was visual.

Because the trip extended late into the night, he retired immediately at its conclusion only to discover that "the movement of the car was retained in kinesthetic imagery and was transferred in perception to the bed upon which I lay." His

⁴ K. M. Dallenbach, Blindfold chess: The single game, *Studies in Psychology, Titchener Commemorative Volume*, 1917, 214-230.

⁵ Warren, *op. cit.*, 455 (italics his).

⁶ Dallenbach, Recurrent images, this JOURNAL, 15, 1924, 155.

travelling companions reported like experiences. Though these kinesthetic after-images had never before been reported, he was of the opinion that they were common, being like those experienced by voyagers when they step ashore after a long trip at sea.

He alone among those taking the automobile trip reported visual after-effects. Regarding them, he wrote:

"I . . . experienced visual after-images of movement which were projected upon the field of my closed eyes or, with open eyes, within the darkness of the bed-room. The movement was toward me, a positive duplication of the perceptions aroused from watching the roadway. During the entire time of driving [he drove the entire distance] I kept my eyes fixedly upon the road directly ahead of the car. . . . Stimulation was constant and of long duration; it was also intense, particularly during the last half of the trip which was made under the illumination of strong headlights."⁷

In 1929, Goodman and Downey published a laboratory study of the images of spectacle rims frequently observed by one of them (Goodman).⁸ Though these authors did not classify Goodman's images, his experiences have all of the earmarks of the recurrent images mentioned above. Goodman described the images and the conditions under which they occurred as follows:

"I have worn glasses for eleven years. . . . I wear them usually all day. . . . The present pair, with heavy black frames, have been worn for a year and a half. When light conditions are right and the glasses are off, an image of about an inch of the bows and the upper outside half and the lower outside quarter of the rims appears."

Writing in the third person, the account continues: "An image of spectacle rims . . . always appears when the spectacles are off and the illumination is reduced to a minimum. . . . This image is so extraordinarily like the perceptual experience that the observer has at times been deceived into thinking he was wearing his glasses when they were off. The failure to confirm his visual experience by a tactual one when he raises his hand gives him a feeling of bewilderment."

Goodman and Downey gave no explanation of these images nor did they attempt to classify them, but they did demonstrate by tests that they were not positive after-images.

If Jaensch's classification of the images into three types:⁹ after-images, eidetic images, and memory images is accepted, the recurrent images described here would have to be classed as eidetic images. They are, however, very different from the eidetic images described by Jaensch. The conditions for their stimulation are not the same. According to

⁷ Dr. Helen M. Richardson, in a personal communication, reported similar experiences after a long drive through winding roads. On going to bed she saw road signs coming toward her which she had carefully observed during the trip.

⁸ G. J. Goodman and J. E. Downey, An image of spectacle rims, this JOURNAL, 41, 1929, 498-502.

⁹ E. R. Jaensch, *Eidetic Imagery*, 1930, 1-30.

Jaensch, the eidetic images can be seen after a brief stimulation. Recurrent images, contrariwise, occur only after long and intense stimulation. Again, such evidence as we have seems to indicate that recurrent images are located in the retina; eidetic images are projected into the environment. Recurrent images tend to occur only with eyes closed or in very dim light; eidetic images can be seen with eyes open in broad daylight.

We have the choice then of breaking up the eidetic image into two types: recurrent images and visualizations; or of limiting eidetic images to visualizations and regarding recurrent images as an additional type. In either case positive images would consist of four types, varying from the usual retinal positive after-image, through recurrent images, visualizations or eidetic images, to the cortical memory images. It would probably be a hopeless task to attempt to divide these phenomena into rigidly independent categories at our present state of knowledge of them.

New Jersey College for Women

NELSON G. HANAWALT

'ATTRACTION' IN FIGURAL AFTER-EFFECTS

Generally speaking, the figural after-effect is characterized by a perceptual shift of some part of the test-figure *away* from the area previously occupied by the inspection-figure; thus, the inspection-figure seems to 'repel' the test-figure.¹ Occasionally, however, the shift is in the opposite direction, and some part of the *T*-figure moves *toward* the region of the *I*-figure; in this case, the *I*-figure seems to 'attract' the *T*-figure.

The attraction-effect is clearly described in Köhler and Wallach's well-known monograph,² as well as in Köhler's earlier *Dynamics in Psychology*.³

¹ Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. Philos. Soc.*, 88, 1944, 269-357, esp. 292.

² *Ibid.*, 296-297. The actual statement is this: "If the *I*-object is an oblong, and if two *T*-squares are shown within the affected area the distance between these squares is shortened. On the other hand, if with the same *I*-object the *T*-objects are shown at a small distance, one above and the other below the oblong, their distance will be greater than the physically equal distance of two *T*-objects in an unaffected zone. Between these two situations the transition is not abrupt. Suppose that the *T*-squares lie half inside and half outside the area of the *I*-oblong. If the inspection period is not too short the distance between these squares will decrease. Even with *T*-squares which lie entirely outside the oblong but just adjacent to it the same will be observed—always provided that the inspection period has been sufficiently long. [Italics added; at this point in the original there appears a cross-reference to Köhler's *Dynamics in Psychology*, 1940, 99, where there is detailed and diagrammed a similar demonstration.] It is true that with an oblong of small width the effect tends to lie in the direction of decreased distance between the *T*-objects when inspection has been sufficiently long. But with an *I*-square of considerable size it will be more difficult to obtain the same result unless the inspection period is very long indeed."

³ Köhler, *op. cit.*, 99.

A recent discussion of the theory of figural after-effects has, nevertheless, brought forth a denial of the existence of the attraction-phenomenon.⁴ It therefore seems appropriate to report some observations which seem to demonstrate the reality of the effect.

The new evidence is much like that presented by Köhler and Wallach. As shown in Figure 1A, the test-figure consisted of two $\frac{1}{8}$ -in. squares (*T*), each $\frac{1}{4}$ in. to the right of the fixation-point; one square lay $\frac{3}{16}$ in. above

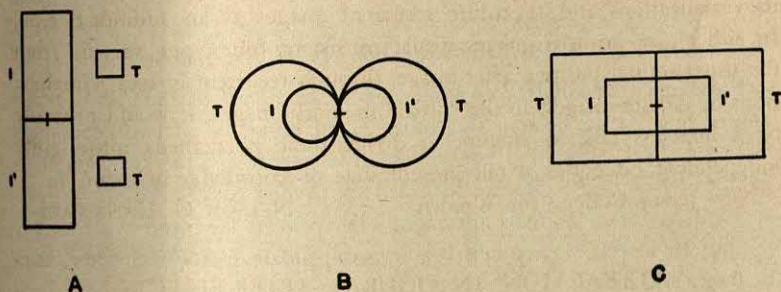


FIG. 1. DEMONSTRATIONS PRODUCING THE ATTRACTION-EFFECT

the level of the fixation-point, the other $\frac{3}{16}$ in. below. The inspection-figure was an oblong, $\frac{1}{2}$ in. vertically by $\frac{1}{4}$ in. horizontally, set either above or below the fixation-point (*I* or *I'*, Fig. 1A). The *T*-figure was always displayed 12 in. from *O*. The *I*-figure was sometimes displayed at 12 in., but usually it was represented by an exactly four-fold enlargement set at 48 in., providing an optically equivalent situation. Inspection was monocular, and of 10-min. duration. A head-rest was provided; illumination was even; and the experimental room was isolated, light-proof, and painted flat black. The *O*s were seven male graduate-students. Six of them reported some after-effect; in all cases it involved movement of the critical *T*-square toward the area occupied by the inspection-figure. Generally the effect was small, but in one instance it was reportedly quite striking.

Somewhat different figures give, at least for the writer, similar movements of *T*-contours toward *I*-contours. Fig. 1B, for instance, shows a *T*-figure consisting of two horizontally tangential circles, the fixation-point being at the point of tangency; using either of the smaller circles (*I* or *I'*)

⁴Charles E. Osgood and Albert W. Heyer, Jr., A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1952, 98-118; Kendon Smith, The statistical theory of the figural after-effect, *ibid.*, 59, 1952, 401-402; Osgood, Kendon Smith's comments on A new interpretation of figural after-effects, *ibid.*, 60, 1953, 211-212.

as an inspection-figure, and employing a 10-min. inspection-period, one sees shrinkage in the enclosing *T* circle. The demonstration is made somewhat more striking if the circles are replaced by squares, as in Fig. 1C; here, one may expect shrinkage in the affected *T*-square—and shrinkage which is quite discrete, so that now the *T*-figure seems to consist of two nonequivalent squares in abutment.

Pennsylvania State University

KENDON SMITH

THE 1953 MEETING OF SECTION I, AAAS

Programs sponsored or co-sponsored by Section I of the AAAS at the annual meeting in Boston, December 27-30, consisted of two general sessions of short, submitted papers and five symposia. One symposium organized by Burton S. Rosner included six invited papers dealing with the social behavior of animals and men. The speakers surveyed evidence pertaining to insects, lower vertebrates, sub-primate mammals, non-human, primates and man.

A second symposium on brain function was arranged by Walter Rosenblyth. The five papers dealt with effects of anesthesia on evoked potentials, relations between the nervous system and the adrenal cortex, functions of the inferior temporal cortex in visual discrimination, and effects of brain injuries in human beings. E. B. Newman organized a symposium of six papers on sensory processes. Topics covered included brightness contrast in pigeons, vibratory sensation in the skin as a basis for communication, a neurophysiological model for olfactory discrimination, scaling procedures for loudness and brightness (two speakers), and a discussion of the interpretation of neurological models for psychophysical phenomena.

A series of four invited papers on human engineering and information theory was arranged by Leonard C. Mead. These papers were concerned with information theory as a model of the organism, as related to visual displays, as related to display-control relationships, and as it pertains to man-machine systems.

In cooperation with Section H (Anthropology), Section I sponsored eleven invited papers treating primate evolution and behavior. The first session consisted of eight papers dealing with evidence concerning evolution of primates, including man. The three papers of the second session dealt with grouping behavior of non-human primates, mental evolution in primates, and the cultural capacity of the chimpanzee.

Together with the Committee for the Study of Animal Societies under

Natural Conditions, the American Zoological Society, and the Ecological Society of America, Section I shared the responsibility for a program of submitted papers dealing with a variety of subjects ranging from factors controlling ovulation in barn swallows to the experimental and hormonal control of sexual behavior in male cats.

The remaining open session of submitted papers also dealt with a number of different subjects including investigations of the Rorschach Test, activity levels of young chimpanzees and human psychotics, emotion and coronary disease, and physiological damage resulting from emotional stress.

The Vice-Presidential Address delivered by Frank A. Beach was entitled, "Development of the individual from conception to conceptualization."

Donald Lindsley was elected Vice-President of the Section for 1954, and J. C. R. Licklider was elected Member-at-Large to the Council, to serve for three years.

Yale University

FRANK A. BEACH

SIXTY-FIRST ANNUAL MEETING OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

The sixty-first annual meeting of the American Psychological Association was held in Cleveland, Ohio, September 4 to 9, 1953. The attendance was over 4,000, and 383 research papers, 76 symposia and discussion groups, plus numerous addresses, business meeting, and various informal gatherings were scheduled. Almost all sessions were sponsored by the APA's seventeen divisions and thus reflected the various specialized interests of psychologists.

The annual presidential address was delivered by Laurance F. Shaffer. Entitled "Of whose reality I cannot doubt," the address reported the results of a study designed to test the hypothesis that psychologists' approaches to psychological phenomena are determined by their perception of reality, and that their perceptions are, in turn, determined by their own experiences.

Gardner Murphy, as recipient of the Kurt Lewin Memorial Award presented by the Society for the Psychological Study of Social Issues—A Division of the APA, spoke on "Human potentialities." In a speech sponsored by the Conference of State Psychological Associations, David W. Louisell, professor of law at the University of Minnesota, discussed "The psychologist in the legal world." Of significance to the development of psychology

as a profession was the adoption, as official Association policy, of a set of principles formulated by the Ad Hoc Committee on Relations with Other Professions.

It was announced that E. Lowell Kelly was elected president-elect for the ensuing year. The 1954 convention will be held in New York City on September 3 to 8.

Washington, D. C.

FILLMORE H. SANFORD

TWELFTH ANNUAL MEETING OF THE CANADIAN PSYCHOLOGICAL ASSOCIATION

The Canadian Psychological Association held its twelfth Annual Meeting at Queen's University, Kingston, Ontario, May 28-30, 1953. Approximately 200 members were present, and 41 scientific papers were read. Three symposia were held on the topics, "Graduate training in psychology," "Comparative psychology," and "Research methodology in social psychology." An invited address was given by Edwin C. Boring of Harvard University on, "Psychology: whence, where, and even whether." The Presidential address was given by D. O. Hebb on the topic, "On human thought." Entertainment was provided through the courtesy of Queen's University.

The officers for the coming year are: Honorary President, William Line; President, D. C. Williams; Past President, D. O. Hebb; President-Elect, Father Noël Mailloux; and Secretary-Treasurer, George A. Ferguson.

The thirteenth Annual Meeting of the Association will be held in Montreal on June 5-6, 1954, immediately prior to the fourteenth International Congress of Psychology which is being held in Montreal, June 7-12, 1954.

McGill University

GEORGE A. FERGUSON

ERRATA

The footnotes on pp. 311 and 312 of the April 1953 issue (Vol. 66, No. 2), both of which are numbered 1, were mistakenly interchanged in make-ready. The footnote on p. 311 belongs on p. 312 and the one on p. 312 on p. 311.

K.M.D.

Harry Miles Johnson: 1885-1953

Harry Miles Johnson was born in Nelson, Missouri, on May 16, 1885, and died in New Orleans on August 16, 1953. He attended what elementary and high schools were available in a rural Missouri district, and in 1909 received his A.B. degree, magna cum laude, from Missouri Valley College at Marshall, Missouri. He entered the graduate school at Johns Hopkins to prepare for the ministry, but shifted to the Department of Psychology where he worked with Knight Dunlap, H. S. Jennings, A. O. Lovejoy, and John B. Watson. He wrote his dissertation, under Watson's direction, on *Audition and Habit-Formation in the Dog* and published it as Behavior Monograph No. 8, in 1913. In a list of his 'teachers' he included Percy Cobb, J. E. Coover, and Harvey Carr, although he had no formal courses with any of them. Some of the graduate students of his day were English Bagby, Buford Johnson, Curt P. Richter, G. R. Wells, and Karl Lashley. He received his Ph.D. in 1912.

In association with Percy Cobb and others, mostly physicists, he worked as psychologist on problems of vision at the Nela Research Laboratory from 1912 to 1918. During these and a few following years he published nine articles on his work in animal vision, six articles on problems of human vision, and contributed a chapter, *The Sense of Vision*, to Watson's second edition of *Psychology from the Standpoint of a Behaviorist*.

From January 1918 until July 1920, during World War I, he was a Lieutenant and Captain in the Sanitary Corps and later the Chief of the Psychology Section in the Air Service Medical Research Laboratory, where his work was primarily concerned with the effects of oxygen deprivation. Scattered through the following years he published five articles on problems related to aviation.

For a year (1920-1921) he was Consultant on factory lighting in the Engineering Department of the B. F. Goodrich Company. He published two articles on reaction-time and one on attention. From 1921 until 1923 he was Professional Lecturer in Psychology at the University of Minnesota, and from 1923 to 1925 he was Assistant Professor of Psychology at Ohio State University.

From 1925 until 1932 Johnson was Senior Fellow and Head of the (Simmons) Investigation of Sleep at Mellon Institute in Pittsburgh. He was deeply interested in this work as an opportunity to apply psycho-

logical principles and methods to concrete problems in industry. During this time he contributed a chapter on sleep to W. L. Valentine's *Readings in Experimental Psychology*, issued one private printing on *Bodily Positions in Restful Sleep*, gave two radio addresses on the subject and published ten articles and four reviews on the subject of sleep. He maintained his interest in this subject and wrote occasional reviews of articles on sleep until 1938.

From 1931 to 1936, overlapping the Simmons investigation by a year, he was Professor and Head of the Department of Psychology and Education at the American University in Washington, D.C., and in these years he taught in three summer sessions at the University of Virginia. During this time he showed a great interest in testing and statistical methods that lasted through the rest of his life for he published eleven articles and reviews on these subjects. From 1936 until 1938 he was Research Associate on the Highway Research Board of the National Research Council, in charge of a study of human causes of highway accidents. With a staff of some 75 persons, accident records of over 29,500 Connecticut automobile drivers were studied. Beside his official reports he wrote four articles, and reviewed two, on accident proneness.

In 1938 Johnson went to Tulane University as Professor and Head of the Department of Psychology and in 1948 was made John Madison Fletcher Research Professor. In 1947 he was awarded an honorary LL.D. degree by his alma mater, Missouri Valley College. Before his coming to Tulane there had never been a psychological laboratory there, and no work at the graduate level was offered in the psychology department. Johnson took great pride in, and derived much pleasure from, the establishment of a laboratory and the initiation of graduate work. The new laboratory was located in an attic, an old story to many of the older generation of psychologists in this country. It has been twice moved and enlarged and still is temporarily housed, but it is now permanently established at Tulane. During his years at Tulane he published twenty articles, about half of which were on tests and statistical methods. During World War II he spent part time for five years on the problems of selecting and training aircraft personnel. He retired from the faculty of Tulane in 1950 at the age of 65 years, but for another year he had an office on the campus, directed a project for the armed forces and did some consulting work.

Dr. Johnson was a member at large of the National Research Council from 1926 until 1929 and from 1937 until 1938, Vice-Chairman of the Division of Anthropology and Psychology in 1937-1938, member of the

Board of Appointments to post-doctoral N.R.C. fellowships in the Biological Sciences in 1930-1933, Chairman of the Committee on the Psychology of the Highway in 1934-1938, executive member of the Committee on the Selection and Training of Aircraft Pilots in 1939-1946. He was a fellow of the A.A.A.S., American Psychological Association, National Institute of Psychology, Society of Experimental Psychologists and the American Association for Applied Psychology. He was a member of the American Statistical Association and the Southern Society for Philosophy and Psychology in 1913-1919 and after 1926, a member of its council in 1930-1933 and in 1936-1937, and its president in 1931-1932. He has been a coöperating editor of this JOURNAL since 1940.

Perhaps his greatest contribution to psychology arose from his close association with individual colleagues and students, his insistence upon clear thinking and a real purpose in work. Although he maintained an interest in many subjects, especially in symbolic logic (in which he had worked out a system before there was an available text on the subject), he never wavered in his interest in experimental psychology. His concern in other subjects was secondary and his knowledge of them was brought to bear upon psychological problems whenever applicable. He was impatient of experimenting just for the sake of experimenting, and he vigorously disapproved hypocrisy, loose thinking and mental laziness. He once said: "My main delight is association with intelligent people; my chief aversion is people who live on buncombe, whether by manufacturing it for others or by using it for food."

Harry Johnson's father was a Missouri farmer for whom he had great admiration and respect as well as affection. He valued highly his father's ability to think straight in spite of his having had little formal education. His mother was deeply religious and it was due to her influence and her ambition for Harry that he entered Johns Hopkins to study for the ministry. He gave it up for psychology after he had entered, but in many ways he remained oddly puritanical and orthodox throughout his life. Some seven or eight years ago he joined the Episcopal Church and became, and remained, an active worker at the cathedral in New Orleans. With the passing of Harry Johnson, psychology loses one of its most colorful advocates, and one of its few remaining strong ties with philosophy.

University of Virginia

JOHN PAUL NAFE

BOOK REVIEWS

Edited by M. E. BITTERMAN, University of Texas

✓ *New Means of Studying Color Blindness and Normal Foveal Color Vision: With Some Results and Their Genetical Implications.* By GORDON L. WALLS and RAVENNA W. MATHEWS. Berkeley and Los Angeles, University of California Publications in Psychology, Vol. 7, No. 1, 1952. Pp. iv, 172.

Under a somewhat misleading title, this heady and contentious dissertation proposes a novel version of the entoptic shadow pattern known as "Maxwell's Spot," replacing macular pigment with a suppositious mosaic of foveal red and blue receptors as the explanatory principle. Ignoring as it does the refinements of technique and the backlog of demonstrated color facts of psychological science, the study falls rather within the confines of optometry. Since, however, its labyrinthine argument turns up points in retinal histology and photochemistry of moment in the simplest color observation and as yet inadequately explored or recognized, it merits at least passing attention, however one assesses the value of its "new means of study."

Among the problematic issues raised are the size and contours of the fovea (the spot of clearest vision), of the rod-free area, and of the macula lutea (the central retinal field impregnated with a yellow pigment variously identified as a lutein, a carotenoid, or a derivative of visual purple, the absorption curve of which is commonly held to account for irregularities in color matches, and the failure of cases of deviate color vision to run true to form). Other points of theoretical and practical interest arising are: the precise color temperature of white light, whether 5000°, 6700°, or 7500° Kelvin; visual acuity under different monochromatic illuminations; the peculiar heredity of minor deviations from normal color vision, and the possible intergrading with each other or with normality of the recognized variants; finally, the size and pattern of Maxwell's spot, briefly visible under short-wave stimulation of the foveal area, and currently assumed to represent the filtering effect of a film of macular pigment, which, by absorbing some of the short-wave radiations, protects the cone receptors in brilliant sunlight and at the same time increases visual acuity (by reducing chromatic aberration).

The investigators cited range from Buzzi (1782) and Soemmering (*circa* 1800), through Müller and Helmholtz, Sachs (1891), Von Kries (1897), Greeff (1900), Gullstrand (1905), Holm (1922), Wald (1945), Miles (1947-49), to Ségel and Hartridge (1951). In the discussion over a minute blue-blind (or tetartanopic) point in the fovea, the names of König, Willmer (1949) and Wright enter. In genetics, Trendelenburg, Pickford and others are quoted. But while the authors figure as industrious collaters of source-literature, their reports are to be read with caution. They have an axe to grind—the ineffectiveness of macular absorption to alter color perception (see pages 21 and 82 where macular pigment is termed "the little mother of all discrepancies" and its indiscriminate use by theorists held up to scorn). The main objective of this study is, in fact, the launching of a substitute proposal, which seeks to explain the phantom Maxwell pattern and other matters in terms of non-

uniform distribution of red, green, and blue receptors in the central retinal area, in both normal and protanopic eyes. Promotion of this idea, suggested doubtless by Walls' studies of the mosaic retinae of lower animals (see *The Vertebrate Eye*, 1942), has a secondary aim—cutting away the ground from under four-color theorists, who, as Walls mistakenly believes, are entirely dependent on yellow pigmentation to solve the enigma of protanopia, its divergence from deuteranopia, and the idiosyncrasies of both types. (The use in theory of macular pigment, by the way, stems not from Hering but from Maxwell.)

Tenacious adherents of the old Thomas Young hypothesis of triune hue receptors (with minor modifications), the authors are set to go any length to sustain it. Masters of argumentative chiaroscuro, they skilfully highlight favorable points in the literature, and shadow others, e.g. Hecht's neutral-point data, and Ahlenstiël's Maxwell-spot studies. Curious gaps appear in the citation of authorities; e.g., in discussing the topography of the fovea (really a peculiar pit) and the inner neural layers of the retina, all reference to Polyak's monumental work of 1941 is omitted, and though the 9-page bibliography (pp. 164-172) lists 124 authors, his name is missing. Further, though the only photo-sensitive or selectively absorbent pigment (apart from the macular) so far demonstrated in the human retina is rhodopsin (with porphyropsin and iodopsin in the infra-human), and though Sheppard's article of 1945 on variable dyes, and Ségal's of 1951, foreshadowing his *Mécanisme de la Vision des Couleurs* of 1953, are listed, photochemical variation of rhodopsin is summarily dismissed as having a possible bearing on deviating types of color vision. Instead, the mythical photochemical pigments of trichromatic tradition—red-, green-, and blue-sensitive—are used, variable in amount only, to sustain the weight of the research in hand, explain both normal and abnormal color perceptions, and the presence or absence of the elusive 'Maxwell's spot.'

Prefatory to the experimental record, 59 pages are devoted to exposition of the *a priori* premises and general frame of reference of the study, along with the unique phraseology and catchwords to be employed—pro-sensation, intensitivity, chromaticness, anomal, the excess, loss, and collapse hypotheses, and the RDP. *Normal color vision*, *Big Four color blindness*, *the excess hypothesis*, *genetics of color blindness*, *Maxwell's Spot* (24 pages), and *receptor type distribution (RDP)* are section captions. The critical area for color perception, we learn, is in the retina, not the cortex, and the cones are pretty surely the media of differentiation. (This emphasis, which may conceal a fling at Müller, ignores Polyak's and Houstoun's suggestions that the synaptic layers of the retina may be the selective ones, as well as Klüver's vote for the geniculate-striate area. In the final effort to dispose of 'peculiar cases,' however, we shall find Walls weakening, in favor of a central as well as a retinal protanopia.)

While cones are red-, green-, and blue-sensitive, each equipped with one of the three selective pigments or color filters, yellow is merely a synthetic quality, like wetness, having no cone of its own. (Here the findings of Hartridge and Motokawa with punctiform or electric stimulation go unnoticed; while Hurvich and Jameson's refutation of Hecht's faulty technique and logic in proclaiming 'binocular fusion' of yellow as cortical, hence a valid stanchion of Helmholtz' theory, is summarily dismissed.) A bluish yellow is inconceivable—though why remains a mystery (except to the opponents or four-color theory). White is neither a synthetic quality (the sum of three hues), nor has it a specific cone, nor is it a summation of the specific

brightnesses of colors. White is simply an excess-phenomenon, existing by default—a notion only mistily elucidated, but useful to account for the persistence of white in dichromatic vision without either coloration or loss of intensity supervening.

The handling of the *Big Four of deviate color vision*, the dichromatic protanope and deuteranope, and the trichromatic protanomalous and deuteranomalous, is in part eclectic, in part *sui generis*. Protanopia, marked by long-wave depression of luminosity and a relatively short-wave neutral point, is a "loss" phenomenon, with the red receptors entirely missing and color perception limited to blues and greens. The O's notion that he sees *yellows* (supported by the testimony of Dr. Pole and mis-diagnosed unilateral cases, e.g. Von Hippel's Goldenberg) is illusory—though on page 87 the protanope's conviction that he sees *blues* is declared irrefutable. Deuteranopia (the green-blindness of Von Kries and the older literature), with the spectral location of its luminosity curve practically normal, and its neutral point nearer 500 m μ , is not a "loss" but a "collapse" phenomenon—an idea first suggested by Fick back in the 1870's, an admission of the fact that 'red' and 'green' blindness are not parallel cases. Briefly, the red and green curves collapse on each other, the two sets of receptors lose their identity, become sensitized to a wider range of wavelengths—and the result is perceived as yellow. This "collapse" might be due to a change in photochemical composition, equivalent to the genesis of a yellow cone, hence reminiscent of Ladd-Franklin's evolutionary view. The latter, however, is declared "zoologically repugnant," and Walls finally settles on the notion that each cone stocks both red- and green-sensitive pigment, unchanged but in varying proportions—a decision useful later in explaining anomalous trichromates and their intergrading with normal-visioned. The hues seen by deuteranopes are of course blue and yellow. White is not reduced in brightness in either type of dichromate, as Hecht's and Hsia's faulty technique would seem to indicate.

In "anomals," saturation is greatly depressed throughout the spectrum, and chromatic adaptation rapid, according to Wieland and Chapanis. The anomaloscope equation—the red-green mixture necessary to match yellow—distinguishes them, a high proportion of green marking the *deuteranomalous*, of red, the *protanomalous*. The first, after much discussion, is classed with deuteranopia, as an example of "collapse," with degrees of "substitution" instead of slight differentiations of pigment as the explanatory principle. As to its intergrading with deuteranopia or normality, opinions differ, but at least two degrees, mild and severe, of deuteranomaly on its way to dichromatism, are accepted by geneticists (p. 13). As to protanomaly, its occurrence in a continuum of grades between normality and dichromatism (protanopia) is probable, according to Trendelenburg. Pitt's views are found inadequate; Pickford's impossible (p. 15). (Incidentally, Walls' chromatic sensitivity curves (p. 2) are admittedly fictitious, showing the red and green curves nearly coinciding in the normal, so that the collapse of the green curve on the red is not difficult to figure—though variable views as to its cause or nature crop up throughout the text. Obviously, the quantitative terms in which the "loss" and "collapse" hypotheses are framed—even while ruling out quantitative changes in the photochemical molecule—favor the admission of intermediate grades between dichromatism and normality, though failing to explain why deuteranomaly will be found to outnumber all other types in the ratio of 5 to 8.)

Turning to the "excess hypothesis" (a view similar to Szekeres' of 1948 and

Willmer's of 1948-49), Walls argues for the whiteness of the dichromat's white, on the testimony of unilateral cases (earlier questioned), and his own belief that "the dichromate, from his scotopic experience, knows what 'white' means to a normal." The existence of a neutral point in dichromatic vision is accepted, explained by the crossing of two hue curves (an idea borrowed from the opponents theory?). Reluctant admission is made that "huedness and brightness are to some extent independently manipulable."

In a 7-page discussion of the *genetics of color blindness*, the framework of genes, phenes, and multiple alleles (with possible mutation) is invoked to account for anomalous cases and their intergrading with normals, and to establish separate loci in the chromosome for deuteranopia and protanopia (the latter point aimed especially at the proposition that 'protanopia' may be only a variant of deuteranopia, both deviants being red-green blind, but with an added scoterythrous factor in the protanopic, producing a shortening in the red, or a shift of the entire visible spectrum toward the short-wave region). Secondly, the older view that color deficiency is strictly sex-linked, and that the daughter of a dichromat is a *heterozygote*, a 'carrier' and potentially the mother of dichromats, but herself normal-visioned, is replaced with Fleischer's discovery in 1920 that she may be color-weak, or an anomal, or show darkening if not dimming in the red.

Maxwell's spot gets the lead in pages 27-51. Here the above-noted lack of agreement among research workers and authorities as to the absorption curve and distribution of macular pigment and the size of the rod-free fovea (readily understandable in view of the difficulty of postmortem examination, and the fact demonstrated by Polyak that the fovea is actually a small crater, with cones gradually slimming and lengthening from margin to centre until they resemble rods) is used to clear the ground for Walls' unique decision later that a 3° area (coterminous with the 'clearing' of his Maxwell pattern) is *pigment-free*, while a $33'$ center (the foveola?) is picked as the rod-free fovea. As for macular pigment, its use by Sloan and Judd to interpret odd unilateral cases (1948), or by Holm (1922) and Miles (1947-49) to interpret decrease of blue in a Maxwell pattern viewed through a violet or purple dichroic filter, is discredited. Varying photochemical concentrations in receptors, or non-uniform distribution of the three type-cones are pronounced superior principles of explanation. Though rejecting Gullstrand's view of macular yellow as a post-mortem artifact, and of the halo and central dot of the Maxwell pattern as purely optical effects, Walls holds pigment concentration in the central thinned-out retinal area to be absent or ineffectual (tossing into the discard his statement of 1942 in *Visual Mechanisms* that in primates the yellow pigment envelops and includes the fovea—a definitely utilitarian device to sharpen vision, akin to the yellow oil drops in the lizard). Curiously, the fact is ignored that any reader can demonstrate for himself, without benefit of apparatus, the macular pigment and Maxwell's spot. Both are visible, when projected on a white ground by a momentarily dark-adapted eye, as a thin yellowish disk with central dot, clearing, and shadowy halo, the whole subtending $3-7^\circ$ by a simple calculation. Similarly, the fovea, mapped by letter-span on the printed page, shows Walls' $33'$ choice (based on old 1907 data) far too small; Polyak's $1^\circ 40'$ finding seems to fit the facts closer.

Receptor-type distribution (pp. 51-59) disposes of Hecht's fruitless 20-year endeavor, through the study of acuity and luminosity, to shore up the trichromatic

theory, support the thesis of equal numbers and uniform distribution of the three type-cones, and read off both deuterans and protans as cases of "loss." Hecht's brief for lowered luminosity in all dichromats is rejected as unsound though his demonstration of subnormal acuity of protans in red light, with compensatory heightening in blue, is accepted (as likely to prove useful later). Hartridge's views on random distribution and 'cluster' grouping of receptors also are questioned, as lacking confirmation.

On the other hand, the König-Willmer-Wright contention for a minute blue-blind foveocentral area, seconded by Stiles' charting of thresholds across the fovea with 10' monochromatic lights, plays directly into Walls' hand. As histologist and geneticist he speaks out for a "rigidly planned," non-random and non-uniform distribution of receptor types, with red cones predominating in the normal fovea and replaced by blue ones in the protanope. (No reason for this partiality to blue in "the central planning committee," usually designated in anthropomorphic terms, is advanced. Nor is there any reference to the long current view of "planned" distribution in the paracentral area, where a red-green-blind (or weak) zone merges into an outer achromatic one.)

The ways are now cleared for the direct observation of the fleeting entoptic phenomenon known for its discoverer as Maxwell's spot, but here rechristened **RDP (receptor distribution pattern)**. A unique red-purple dichromic Wratten filter, No. 2389 (recommended by Miles), the absorption curve for which is described only, is chosen in place of the blue or violet light of earlier observers. All but one of the 64 color-normal Os (14 of whom were women) reported the phantom pattern as some variant of a reddish smudge or cloud, with or without a central redder or darker dot, in a 'clearing' surrounded by a halo, the whole subtending 3 or more degrees on the retina. Variants of the pattern were classified into Types I-VI (surmised as possibly running in families), with 17 showing no differentiation of dot and clearing. Discarding all earlier interpretations, Walls deduces a high concentration of red receptors in the center of the retina, and turns to the protanopic or red-blind, as presumably lacking all receptors of this type, for negative confirmation.

To single out protans (Farnsworth's term) for this purpose, three measures were adopted: (1) the ratio of red to green required to match yellow in a filter anomaloscope; (2) the luminosity (in terms of a voltage-reading only) of red and of blue filtered light in relation to a standard green; and (3) the wavelength of the neutral point as calculated from the ratio of Munsell blue to green needed to match gray on a rotating disk. Results and tables of scores are incompletely or inadequately presented; but though the arrays are practically continuous (except for red luminosity scores) the group of 35 anormals who volunteered was divided, quite satisfactorily to the experimenters, into 19 protanopes and 16 deuteranopes (a reversal of the ratio usual in non-selected groups). Thirteen of the protans reported a blue **RDP**, while only one of the deuterans saw anything on the screen at all. Walls accepts this outcome as proof that the red receptors, overnumerous in the normal fovea, are replaced in the protan by blue ones, which simply crowd in or are directed by an embryonic "planning" agency. Absence of a deutan **RDP** is juggled into proof that this defect is a "collapse" phenomenon, leaving no guiding clues to the embryonic "planner." Failure of the deutan to report a patterned yellow smudge may be due to viewing conditions and the mazda source. Neither here nor in the earlier **RDP** observations were precautions taken apparently as to neutral preadaptation, or

preliminary training in color naming, though yellowish light is commonly confused with white, and violet and purple are terms notoriously interchanged by the uninitiate. Nor was any guard set (except in one test) against chromatic aberration, which overemphasizes red in the near-sighted, blue in the far-sighted.)

Deuteranomalous and protanomalous cases, distinguished from dichromats by lack of a neutral point, are next discussed (pp. 95 to 106) with a number of pedigrees, in an intricate argument that works up into family trees, and is reminiscent of a Freudian attempt to explain the normal in terms of the abnormal. While certain unclassifiable cases give trouble, Walls is satisfied that he has proved conclusively that deuteranopia and protanopia are distinct entities (p. 94); that the latter is not deuteranopia "plus a darkened red factor," as Murray, Pickford, and others have suggested—though a "dimmed red factor," sometimes occurring independently of other defects, is reluctantly admitted. (Mistakenly, the authors assume that this reviewer's scoterythrous concept derives from Hering; it originated with Rivers.)

The discovery (matching Schmidt's of 1934) of heterozygotes (women carriers with dichromatic sons) with depressed luminosity of red as their only deviate feature, and of other *O*s that "defy classification" leads to speculation as to the likelihood of a brightness cone, or of bifurcated paths, and to a limited recasting of premises, such as "collapse protanopia" and "loss deuteranopia." Unquestionably, the authors are to be commended for reliance not on one but on a team of tests (anomaloscope, luminosity, and neutral point, not always, to be sure, wisely set up or administered) for diagnosis. Credit must be allotted also for detailed reports of the pedigrees of certain *O*s and related cases in the genetic literature—a line of work carried on more extensively abroad by Pickford and others. But a genetics based on faulty or inadequate anomaloscope readings (see Farnsworth's recent study) is necessarily hidebound and misleading. "There's many a slip twixt the gene and the phene," is Walls' own conclusion, after a 50-page discussion of unique and puzzling cases. Many of the dilemmas would be readily resolved, however, if the four-color proponents' postulates were to be substituted for trichromatic. To transcribe his comments on another worker: "the three-component theory will be destroyed by inbreeding if it is tested further by considerations that derive from a system that assumes the theory." In the Introduction, however, interpretation on the basis of other assumptions is generously invited.

In the final pages (156-63) the postulates set up in the preamble are amplified or contracted (with some concessions to the obstinacy of observed fact) into 61 one-to-three-sentence maxims. No index is provided, however, which puts the reader at a disadvantage in assessing the frequent textual contradictions, relating especially to the "collapse" theory of deuteranopia, the troublesome "darkened" or "dimmed red" factor, and the significance of variations in the *RDP*. As for the latter, the variant contours reported suggest, to the unbiased reader, not so much peculiar groupings of displaced red or blue receptors, as a mild yellow inundation of an irregular coastline—that of the inner retinal surface in the foveal basin. As for coloring, no one but Walls has ever insisted that the macular pigment, to be visually effective, must intercept *all* the short-wave "pro-blue" stimuli. The residue passed would account for the violet or purplish tones reported now and then in the *RDP* descriptions. The shadowy portions of the pattern may represent, not regimented red receptors, but simply thicker or more opaque tissue, or steep slopes.

As for "a world of work remaining to be done," all research workers in color

must agree. Some, however, may remain sceptical as to the wisdom of concentrating, at the histologist's behest, on the mosaic pattern of 3-type cone receptors in the fovea; sceptical also as to his permanent liquidation of the macula. The failure of certain deuterans to report a definite *RDP* pattern or color may, it is true, be an artifact of method or instructions worth investigating; but instead of the cumbersome set-up with filter No. 2389 to separate deuterans from protans one suggests trial first of the old duplex diagnostic Ishihara scarlet and purple digit-plate. Understanding that in the deutan's perception of the purple digit the blue primary in the mixture cancels the yellow valence of the red, leaving residual grays, while in the perception of the scoterythrous protan, the entire red response including its yellow valence, is eliminated, leaving a blue (intensified by the possible shift of the entire spectrum toward violet), should serve to explain not only why the purple digit on its gray ground is invisible to the deutan and seen as bluish by the protan, but also why the latter sees the *RDP* for Walls in blues and the deutan probably in yellowish tones hardly visible against a mazda-lit ground.

As for useful filters, in view of Ahlenstiel's claim that mazda illumination converts his perceptions from protanopic to deuteranopic, trial of a (partially) minus-red-filter (if such exists) to convert deutan into protan (scoffingly proposed by Walls) would seem to offer a rewarding line of research (possibly already tapped by Collins).

After grappling with the color reports of the *RDP* in this study and with the experimenters' deductions, certain primary qualifications essential to any observer or experimenter, engaging to explore the color-mediating zones of the retina, suggest themselves. First of all, acquaintance with the permutations and commutations of hue, a few hundreds of which may be sampled in the small Munsell handbook of color, is desirable, plus a stable descriptive system, taking some note of the common confusion of brightness and saturation, and the almost universal interchange of the terms violet and purple in the uninitiate. To these should be added familiarity with liminal colored lights and the ease of confusion of blues and greens; knowledge of the dissection of violets and purples in the near- and far-sighted eye into reds and blues, with emphasis on one or the other; and experience of the subtle transformations effected by change in illumination, by adaptation, by complementary contrast and negative after-images, and by the operation of the Bezold-Brücke law at different levels of intensity of monochromatic lights (due to the spread across the spectrum of the response curve of any primary).

Finally, though one may criticize the methods and conclusions of this monograph, its lack of organization of text and tabular matter, and its bias, its authors merit credit for highlighting neutral-point data; for opening up the field of genetics, neglected this side of the water; for exposing as well the need of further detailed postmortem study of the retinae of normal and color-deviating *O*s, with an eye out for the possible presence of rhodopsin in rodlike foveal cones. It should do service also as an eye-opener to the physicist or physician who accepts the old Young-Helmholtz trichromatic theory uncritically, on faith. For Walls, following Hecht's lead, shows a keen eye for the loopholes and weak points in the latter, and, voluntarily or not, has demonstrated how inadequate are its postulates when put practically to work in the no-man's-land of deviating color vision.

Cornell University

ELSIE MURRAY

Individual Differences in Colour Vision. By R. W. PICKFORD. London, Routledge and Kegan Paul Ltd., 1951. Pp. xviii, 386.

This book represents in the main the results of testing more than 1100 individuals for color blindness. Even this large number of cases is regarded as only preliminary to a more detailed and adequate investigation since, as Pickford points out, not less than 10,000 men and boys and 10,000 women and girls would be required to reveal reliably the main trends of color vision in a country the size of Scotland, the country where the tests were made. Of the 1100 Ss, 155 men and 23 women were found to be major red-green defectives. This high proportion is undoubtedly due to the fact that some of the defectives were invited to take the tests because they were known to be defectives. The tests of color vision were made by means of rotating discs and with a variety of color filters yielding results similar to those obtained with an anomaloscope or colorimeter. The author investigated the ability of the Ss not only to name colors and differences between colors but also to make hue and brightness matches. The colors investigated included both the usual psychological primary or unitary hues and intermediate hues such as blue-green and orange. Color defect was assessed in terms of median or mean settings, range of settings, and deviation of settings from normal settings.

The subject of color deficiency is not limited to mere determination of deviations from normal color matching and color naming. The author succeeds admirably in showing that the problem of color weakness and color blindness is practically co-terminus with all problems of vision. Among the many topics treated at more or less length in the book are questions concerning color vision of children, color vision in the animal kingdom, inheritance of color defects, race and sex differences in color vision, primary colors, adequacy of various color vision tests, and theories of color vision, particularly the trichromatic and Hering theories. Unfortunately, however, anyone hoping to gain a clearer or simpler schema or a clarification of the main facts of deviant color vision will not be helped very much by this book. While the usual classification of color deviant individuals into color blind (with subdivisions) and color anomalous (with subdivisions) is followed here, the organization of the book is not such as to leave the reader with a clear picture of the author's extensive findings. The exposition is unnecessarily complicated because the author presents his material on the basis of the tests employed. This division results in interminable repetition since the various types of color defect are discussed in terms of rotating-disk data, paired color-filter tests, brightness, spectrum and micro-color filter tests, the nine-color experiment, the matching of intermediate colors, and, finally, the four-color test. Since each test involves not only broad differences among different types of color deviant individuals but also minor differences in color response due to the type of test employed, it is difficult, if not impossible, to tell what is ascribable to the method of testing and what may be ascribed to the particular type of defect under study. This criticism is not meant to detract from the competence of the author, who before the appearance of this book had already established himself as one of the leading workers in the field of defective vision through his excellent contributions to the various scientific journals. Rather it is meant to point to the real usefulness and contribution of the book: researchers interested in the minutiae of color deviation as revealed by particular testing devices will find the book a mine of detailed information.

It is to be regretted that the experimental results were not presented at least in part in fundamental terms, *i.e.* in the terms afforded by the C. I. E. trilinear mixture diagram or some derivation of this standard method of specification. Once the x , y coordinates of a colored stimulus are given it is completely specified as a color no matter what its spectral composition may be. All stimuli which plot at the same point in the color mixture diagram look identical to the eye and are absolutely equivalent as visual stimuli. Results expressed in these fundamental terms can be related to many other types of visual measurement and can be reproduced. Results of disk-mixtures made under light sources whose spectral energy distributions are not given cannot be compared and interpreted by other workers with the same degree of confidence that is possible if fundamental specifications of the stimuli are given. Color-names like "fawn" and "peacock" which are often used to specify stimuli do not tell exactly what stimulus is meant. Terms such as these were long ago given up in scientific writing about color.

In common with many psychologists the author of this book argues for the primacy of yellow and in favor of a four-primary theory of color vision. In view of recent findings by Granit and others pointing toward at least seven fundamental types of nerve-response to differences in wave length, it looks as if discarding the modified Young-Helmholtz trichromatic theory will make it necessary to assume more than four primary processes to explain color vision. So far as quantitative considerations apply, it doesn't matter how many independent receptors there may ultimately prove to be; if their action can be accounted for mathematically by means of some minimum number of primaries, this minimum will probably be used. At present it looks as if a modified C. I. E. standard observer employing three primaries will serve as the basis of color measurement for a long time to come.

University of Texas

HARRY HELSON

Anatomy of Public Opinion. By JOHN NORMAN POWELL. New York, Prentice-Hall, 1952. Pp. xi, 619.

Does the current political picture fill you with gloom? Are you apprehensive over current Communist propaganda offensives? Or ditto for Capitalistic propaganda? Do you perchance feel that the contemporary vehicles of communication deluge you with moral judgments disguised as insights and understanding of social phenomena? Does what our citizenry say and do seem so incredibly naïve or foolish, while what they ought to do is so obvious? Are you contemptuous of what the 'so-called' social sciences have to offer us in the way of pearls of wisdom? Are you appalled at the increasing powers of the state? Are you the bedeviled instructor assigned to teach a course in 'Public Opinion' or 'Propaganda'? Well then, here is a volume made to your measure.

It is now some thirty years since the appearance of Walter Lippmann's *Public Opinion*, a volume still memorable for its shrewd substitution of a psychological determinism for the naïve notions that have dogged social theorists and moral philosophers from the beginning. Freshly before Lippmann at that time was the record of the manipulation of attitudes and opinions on a previously unheard of scale that was a concomitant of World War I. There the ideals and symbols that integrate and activate collective action had to be attenuated and revised to fit the needs of unprecedented situations. Where ideologies formerly had only to encompass and alibi

the activities of a single national unit, they now had to transcend national and class lines in an unprecedented manner, if the Allied cause was to win. What with the turn that events took in Russia, there resulted an era of soul-searching scepticism that may well have been disastrous to national mythologies, but it certainly did usher in a period of critical realism in social theory. Significantly enough, Lippmann foresaw the rôle to be played by an emergent, quantitative, social science. Powell's volume does an excellent job of outlining just what the various social sciences have contributed in this area since then. In the formidable, and excellent, appendix devoted to references there is hardly to be found anything antedating 1925!

In brief, here is what the author provides us: First, a brief chapter outlining the problem: definitions, semantics, and a few pages on 'power' and 'freedom' that should be required reading for all, but particularly for those spokesmen for the vested interests currently embattled under the banner of 'freedom.' The second and third chapters are devoted to a brief but adequate summary of methods and problems in opinion-measurement, not so detailed as to be boring to any but a specialist, but sufficiently so to give any intelligent reader a grasp of the problems encountered in opinion-measurement, the pitfalls, the limits of accuracy, the incredible fluidity of the thing measured. There is also a good indication of the kinds of attitude that have been measured and to what degree of success. Then there is a frank and insightful discussion of the social consequences of polls and polling, which is illuminated, as always in this volume, by such evidence and research data as are relevant. The author's evaluations of current work seem to this reviewer to be excellent, but the reader is spared much of the detail concerning the techniques of research used. Polls and polling come in for a major share of the attention, but questionnaire-studies, tests, projective techniques, depth-interviews, mass-observation, and content-analyses are not overlooked. Even the values and evils attributed to the measurement of public opinion are summed up for the reader, all in a manner that might well be taken as a model of objectivity.

The concrete instruments of public opinion and propaganda then come under discussion and analysis, the fourth chapter considering groups and individuals. Here we have a realistic analysis of the actual power-residuals in a society such as ours. Far from tending to see social power as evaporating into abstractions and fictions, the author pin-points individuals at points of control. In this chapter we are given an excellent summary of such studies as have been made on the varied rôles and personalities involved in leadership. Leaders are important precisely because they really do exercise control. The variability in the findings with regard to leadership is well reviewed. The importance of organized groupings in practical and effective politics is effectively set forth, along with a sharp reminder, to those who are contemptuous of 'politics,' that group existence makes that sort of activity unavoidable. The author shows that the unit of organized group life is the pressure-group, and in some detail he pictures the effective pressure-groups now constituting democratic society as illustrated in these United States. The pressure-group gets results by 'lobbying.' Such registered 'pressure-groups' at the national capital numbered well over a thousand in 1949, and their total activities are considered by some to be a billion-dollar industry. As an example, organized medicine mobilized a three-million-dollar war chest to fight socialized medicine. The changed character of the pressure-group in the Soviet Union is briefly commented upon; significant is the fact that pressure-

groups are not eliminated there, they merely take a different form. Yet the author makes a good case for his thesis that the various pressure groups are the basic roots of the democratic process, which, realistically seen, is based on competition between pressure groups. The chapter also includes the suggestions of the American Political Science Association's committee for improving political processes, along with a general summary of research findings concerning public opinion on key current issues.

The press, motion pictures, radio, and television each rate an entire chapter and it is interesting to note that in 1921 only the first of these rated separate-chapter status in Lippmann's work. At that time, motion pictures rated five scattered references in the whole volume; the other two were hardly on the horizon. The chapter in the current volume on the press is a model of calm objectivity. Well, for one thing, the press is considerably reformed in its worst abuses since the days of Sinclair's *Brass Check*, and for another the readers have over the years developed a certain immunity to the control by direct suggestion once exercised by the press. The author uses his pages wisely in reporting the varied factual findings that have something like statistical respectability and throw however feeble a light on the actual status of the press as an opinion making force. The reader is left to draw his own conclusions.

Two chapters are devoted to such experimental evidence as is available on the determinants of opinions and attitudes, the effectiveness of various kinds of appeal under varying conditions. These chapters do an excellent job of summing up such tentative conclusions as might be warranted in a field of almost unbelievably frustrating complexity. There is material here for the professional propagandist, as well as for the individual who wants to keep himself immunized against the next high-powered propaganda offensive. The significant findings in the field seem to be well reported.

Finally, we have a chapter on the rôle of public opinion in a Democracy, which is about as insightful an analysis of the limitations as well as the virtues of the democratic process as this reviewer has seen. Certain it is that improvements in the collective living to which man is obviously doomed will stem from a sound knowledge of the actual processes of social relations that volumes such as this so intelligently organize and provide. Dr. Powell has made somewhat less hopeless the prospects for the appearance of specialists in the field of the social sciences to which we may turn as confidently as to scientists in the older fields, whose subject matter is somewhat less charged with emotionality than are almost all areas of social behavior.

New York University

GEORGE B. VETTER

Modern Science and Modern Man. By JAMES B. CONANT. New York, Columbia University Press, 1952. Pp. 111.

There is little specifically about psychology in this attractive, deadly serious, attention-compelling volume of but four lectures; yet psychologists will read it because they are interested in the scientific nature of their profession in general and in the relation of scientific fact to social value in particular. Let me attempt to extract the gist of Conant's timely argument.

Science is to be regarded as a *limited* activity. For three centuries it has had a continuously increasing success, and what has begun so auspiciously will doubtless continue into the future; yet it would be wrong to suppose that science can ever

exhaust knowledge or that its methods—and Conant insists that there is no one scientific method but many procedures and devices—can ever be adequate to all knowledge or to the whole of reality. Conant appeals here both to the transitory nature of scientific truth (phlogiston has gone and so has the luminiferous ether) and also to the complementarities (not only Heisenberg's uncertainty, but also the necessity of believing simultaneously in incompatibles, like the wave and quantum theories of light, both of which are true while each contradicts the other). It is this twentieth-century revolution in physics that convinces scientists that there are unknowable regions of knowledge, or, if one follows operational logic, that these regions, being unknowable, can not exist in any meaningful sense of the concept of existence.

The notion that science is exploring the total field of knowledge and extending the boundaries of *terra cognita* is all wrong, says Conant. Science is an *activity*. It begins in empiricism, that is to say in the trial-and-error working out of ways to get to results (cooking is almost entirely empirical in this sense) and proceeds toward the building up of theories or models, which have in them all the immutability of logic as well as its tautology. (Surveying is almost all logical theory because it is little more than geometry, mathematics, and simple well-established optics.) Thus the development of a science is shown by the degree of this kind of empiricism to be found in it, and as the degree of empiricism decreases the quantity of theory increases. Physics is far along, as its 'revolution' shows; psychology is just beginning.

Empirical trial-and-error is ever so much less efficient than deduction from established theory. The release of atomic energy showed that and convinced the public too that scientific theory is not a vapor of the ivory tower but a power to be considered in anticipating human weal and woe. Not soon again will the scientist be regarded as an impracticable 'long-hair' as he was before 1940. Indeed, the danger is the other way, that fanatical public belief in science and governmental support of it will lead to foolish expenditure of time and money before empiricism has yielded far enough to theory.

In 1907 J. J. Thomson had suggested that a scientific theory is not a creed but a *policy*, and Conant adopts this view. Science can be regarded as a queer mixture of practical common-sense empirical procedures, general speculative ideas, and abstract reasoning. As such it is an activity, and both its empirical rules of thumb and its generalized theories are guides to action, its policies. These rules are one of the ways in which man in his complex civilization gets along in the world, but they are only one.

Never do you approach a scientific problem without bringing to it many values implicit in your habits of thought, values usually not recognized. Some of these articles of faith are so deeply ingrained that it is almost impossible for man to divest himself of them, attitudes like the belief in a spatially three-dimensional world or in the existence of other persons than the thinker himself. It would be folly to think that science ever gets anywhere without the use of arbitrary and adventitious values, and that is why, as an activity, it is a policy and not a sure road-map to truth.

You have only to extend this comprehension of the relation of science to values to see where spiritual values enter in. Conant suggests that the minimum of additional values needed to specify practical activity as *spiritual* are (1) a belief in a distinction between men and animals (e.g. dislike of cannibalism), and (2) a belief

that personal action is related to some larger whole. For a religious faith you must add still more *a priori*. At these levels you must recognize the existence of various universes of inquiry, controlled by different policies, not by different creeds. Always, always you have the values, but they are different in different faiths, and within science they are different in the consideration of the problem of the origin of life by the cosmologist, the biologist and the chemist, and they are different when light is considered by the physicist believing for the nonce in wave theory and by the physicist, perhaps the same one a little later, believing in quantum theory.

Beyond science you have all of man's other activities, which may be good and are not to be denied. There science's great theories take their place alongside of other great monuments of man's creativity, gravitation alongside of the Parthenon.

The view that Conant discards is *materialism* and thus modern dialectical materialism, the notion that science can embrace all knowledge, that it is a creed that can supplant religion. In a way this volume is the antithesis of that contemporaneous defense of determinism in human behavior, Skinner's *Science and Human Behavior* (1953). Skinner, in claiming for science an exclusive right to understand and control human nature, utters the materialist's challenge to an incredulous world. Conant, I think, answers him.

Harvard University

EDWIN G. BORING

The Psychology of Thinking. By W. EDGAR VINACKE. New York, McGraw-Hill, 1952. Pp. xiii, 392.

"This book," one reads in the Preface, "represents an attempt to survey in a systematic fashion the variety of human thought processes, together with what the psychologist has found out about them. The author has not been guided by a preference for any particular species of theory." The result is a painstaking review of some 600-700 items of study, mostly experimental, set forth in 17 chapters. In the Introduction we are told that "In general, thinking serves in the adjustment of the organism to its environment, both internal and external." This implies a "motor theory" for which there is, we are assured later on, "an impressive weight of evidence" (p. 69). After sketching the Historical Background, subsequent chapters, without any very apparent order, treat of such familiar psychological topics as Consciousness and Attention; Ideas, Imagery, and Imageless Thought; The Mechanism of Thinking; Logic; Concepts; Transfer of Training; Problem Solving; Imagination; Autistic Thinking; Creative Thinking; "The Internalization of Experience"; Personalized Thinking and Attitudes.

With thinking in the service of adjustment, a behavioristic slant is given to the discussion, but no very certain conclusions are drawn. "If the nature and function of the receptors, connectors and effectors are explained, consciousness will also have been in large part explained" (p. 21). But how? Is it "reasonably clear now to most psychologists that the 'mind' and the 'body' are one and the same but that one may describe the phenomena at different levels of abstraction" (p. 42)? Along with such evidence of loose or incomplete consideration, one also encounters significant *aperçus*, such as one on "memory which, instead of being a kind of specific record 'written' on specific neurons, may very well be the establishment of a more or less intricate relationship within the pattern of neurons, perhaps, in the case of a complex memory, scattered over the entire cortex" (p. 60).

Although the author stresses 'dynamics,' he does not seem to have a precise meaning for the operational *relations* to which he so often refers in describing the adjustments which thinking serves. From his experimental sources he draws the conclusion that relationships are established gradually. Though he recognizes that they "are more significant than specific content, or specific operations, or specific rules" (p. 178), he seems to regard them as *gradients* achieved by frequency of association. In addition to gradual analysis and 'trial and error,' *insight* is accepted as being a third "mode of attack" upon problems to be solved. The characteristic 'suddenness' of insight is discounted. Although a reference is given to Lacey and Dallenbach on "Acquisition by Children of the Cause-Effect Relationship" (this JOURNAL, 52, 1939, 103), the more striking demonstration of sudden and lasting insight by Kreezer and Dallenbach ("Learning the Relation of Opposition," this JOURNAL, 41, 1929, 432) is not listed. It is proper enough to be cautious in dealing with sudden revelation, but if one uses the term *insight* one should regard its province in psychology and not destroy the special meaning which Köhler gave it in describing the behavior of his apes.

One is troubled by the suggestion that "at least in its earliest stages, the learning of symbolic responses can be explained as simple conditioning by the repeated linkage of words and objects" (p. 272); for, one asks, how was the original conditioner conditioned? It is also rather naïve to state that "a motive could be a system of visual-thalamic-visceral-cortical-muscular relationships. A concept could be a system of sensory-thalamic-visceral-cortical-muscular relationships. In fact the conception of organization proposed above is quite compatible with the motor theory of thinking" (p. 322). What are these "systems" so confidently summoned to provide a motor for thought? Have they any other meaning than that of a congeries of anatomical juxtapositions?

We may, however, thank the author for bringing together a vast amount of material—chiefly experimental—which he sets forth clearly and tactfully, albeit uncritically. Since without a point of view one can hardly be critical, we must be satisfied with his own modest evaluation of his labors. "Throughout the preceding pages," he writes in conclusion, "we have made a concerted effort to bring together materials of diverse kinds, with special emphasis upon scientific research. We sought to organize these data and to give them theoretical coherence. In the course of our treatment, many more questions were raised than we were able to answer. It is our hope that future investigators can answer them" (p. 362).

Cornell University

R. M. OGDEN

Personality Measurement. By LEONARD W. FERGUSON. New York, McGraw-Hill, 1952. Pp. xv, 457.

We find the stated purpose of this volume in its last paragraph, which reads, "The purpose of this text was to explain several of the major methodologies used in personality-test construction." It is actually not a systematic treatment of basic methodologies, as such. Methodology is emphasized, but it is generally presented incident to the description of standard tests or other devices and procedures. The methodology treated also goes beyond test-construction to include rating-procedures, projective techniques, and the observational and experimental study of personality. Interview-methods are not treated, evidently by intention.

The term 'personality' is not defined. This term is evidently conceived in the usual, somewhat narrow sense that excludes all abilities. It is curious that in American psychology the term 'personality' when applied to an individual usually implies the nature of that individual and refers to the kind of person he is. How one can cover the nature of a person without mentioning his abilities is hard to see. Yet in textbooks of psychology, with few exceptions, intelligence and other abilities are excluded from the term. We even find 'personality' often illogically set over against 'intelligence.' Perhaps we need to adopt the British practice of referring to the non-ability aspect of personality as the *orectic* side. It is the *orectic* side of personality that is treated in Ferguson's book.

The term 'measurement' also lacks definition in Ferguson's book. From the great variety of procedures that he treats under this title one might well conclude that the term has a meaning about as broad as 'description.' Such an extension of a term is an author's privilege, but his readers would probably like to know where he conceives the boundaries of measurement to lie.

Ferguson is strongest when he is describing a personality test and telling how it was developed. He writes with the assurance of one who himself has developed more than one type of instrument and who has used such instruments in practice. The orientation is always toward the practical use of tests and other instruments. Nowhere else will one find in one place descriptions of so many types of device. The listing of standard, published tests is selective, as it should be and as it must be. The selection seems representative, though some readers will probably have different opinions at specific places.

While the description of techniques of test-development is usually full, this reviewer should like to have seen more on general theory of personality-measurement. He is also painfully aware of the many statistical problems in testing theory and in the general theory of psychological measurement that were glossed over. True, there is not space in a book of this character for so much background material, but then the question arises whether students should be introduced at all to the standard tests and other instruments so extensively at the superficial levels before they have been sufficiently grounded in theory. Students may thus be encouraged to build their own tests because it looks so easy. They are thus likely to perpetuate the errors that past generations of test-users have made. Ferguson has offered criticisms of many of the methods, and some of the criticisms are discerning and uncommon, but he has missed many other opportunities to evaluate procedures. It will be up to the instructor who uses this book as a text to supply the student with more background and with precautions concerning the training and the skills necessary to construct and to use devices for measuring personality.

We have here, then, a book directed to the description of tests and other devices for the assessment of personality, their development, and their applications. The descriptions are the competent work of one who has had first-hand experience with that concerning which he writes. The content is long on descriptions of techniques but short on basic theory either of personality or of measurement. The subject has been treated at a relatively superficial level, presumably with the intention of bringing it within the reach of the student who probably is not prepared to use wisely the information he is given.

University of Southern California.

J. P. GUILFORD

✓ *The Psychology of Teaching Reading.* By IRVING H. ANDERSON and WALTER F. DEARBORN. New York, Ronald Press, 1952. Pp. x, 382.

A major emphasis in the present text, which was designed primarily for professional courses on the psychology and teaching of reading, is on the psychological foundations of method. Stress is also placed upon the developmental aspects of reading or the growing into reading, without neglect of the learning processes of the child with which the teacher is necessarily concerned. Some will miss a systematic treatment of typography, readability, and illumination. Although these areas are legitimate parts of a psychology of reading, they are avowedly omitted in order to permit concentration on the psychological facts which apply more specifically to instruction in reading. Major topics considered include reading as growth, reading-readiness, eye-movements, word-perception, learning to read, methods of teaching reading, and appraisal of progress in reading. There is an extensive bibliography.

In the section on "Reading as Growth," the authors present the most complete discussion we have of organismic age in its relation to progress in reading. Organismic age represents the average of growth ages (height, weight, dental, carpal, grip, mental, reading) from one point in time. Cases are cited to support the view that, in certain instances, reading-achievement is a function not of mental age alone but of total growth. Nevertheless, in another place (p. 82) it is recognized that mental age is probably the most important single factor making for success in learning to read, but the practice of defining reading-disability in terms of a discrepancy between mental age and reading-age is questioned. Cases are cited in support of the view that certain children, who are retarded in learning to read, will overcome this apparent retardation without remedial training when total development reaches a certain level. No claim is made that this holds for all cases of reading-disability. The rôle of such physical growth-factors as weight, height, and dentition in relation to reading-readiness will seem overemphasized to some.

The discussion of reading-readiness presents views that are in harmony with the practices advocated by sound writers in the field. One can agree that any remedial program which attempts to force the child to read before he is ready has no basis for existence, but a satisfactory remedial program does not do this. The reviewer considers that evaluation and promotion of readiness must be an integral part of diagnosis and remediation. Presumably the authors have this view in mind.

The studies on eye-movements selected for survey are well chosen in terms of the authors' purpose. The emphases upon the inadequacies of certain early studies, the need of longitudinal studies, and false assumptions underlying pacing are sound. The discussion of word-perception is very well done. Results of recent and the older studies are nicely coördinated. Stress is placed upon those aspects of word perception which are basic in learning to read. In dealing with methods of teaching reading, the authors appraise alphabet, phonetic, word, phrase, sentence and story methods, and they conclude that reading is best taught by a combination of methods. The organization of such a plan is outlined in which stress is placed upon the experience-reading approach. In discussing the evaluation of achievement, critical attention is given to teachers' appraisals and the use of a variety of standardized tests.

For a long time there has been need of a treatise on the fundamental aspects of the psychology of reading. Much of the material in the present text is relevant to that need. Obviously the authors aimed to give enough of the fundamentals to form

a basis for a sound approach to the teaching of reading and they have done this job well. The book should receive a cordial welcome from workers in the field.

University of Minnesota

MILES A. TINKER

Social Science and Psychotherapy for Children. By OTTO POLLAK and COLLABORATORS. New York, Russell Sage Foundation, 1952. Pp. 242.

In this volume an attempt is made to bring scientific principles from sociology, cultural anthropology, social psychology, and the psychology of learning to bear upon a psychoanalytically oriented child-guidance program. It grew out of an in-service staff-training program. To combat an over-specialized viewpoint in a particular agency (in this case, a strong orientation toward individual diagnosis), Dr. Pollak decided to utilize broad social concepts with which the child-therapists had not been familiar. These concepts include the family of orientation (all persons living in the household as distinct from biological parents), social interaction, socialization, cultural relativity and cultural contact, social rôles (age and sex) in youth culture, and reinforcement in stimulus-response learning theory. Pollak intentionally avoided results of statistical observation and the statistical generalizations of social science because such material "might have created reluctance on the part of a considerable number of workers to find use for it (p. 23)." This reviewer regrets that a large body of useful, factual material pertaining to child behavior was thus excluded from presentation to these workers with children. The concepts they studied are undoubtedly useful but by no means reflect adequately the relevant research in child development, which even contradicts some of the generalizations presented in this book.

Because space prohibits extensive discussion of all the concepts, we will mention only three—extra-familial influences in pathogenesis (a chapter written by S. R. Slavson), age and sex rôles in youth culture, and reinforcement in stimulus-response learning theory. The consideration of extra-familial influences is actually a discussion of socialization. Lewin's concept of the life space and the concept of the individual's phenomenal field are introduced. The individual develops as a self only by adapting to various groups. The significance of skill in play or in social behavior for the process of adjustment is not presented directly. The importance of success in the one-sex group so typical of middle and older children is stressed, but success appears to be something awarded arbitrarily by the group rather than recognition accorded to achievement. The various extra-familial relationships are considered important in therapy as well as in diagnosis, and case-material is presented to confirm both viewpoints.

The chapter on age-sex rôles and psychotherapy of adolescents assumes that when an adolescent takes on a rôle not befitting his age and sex category he may be expressing internal conflict. Such a rôle may further complicate the child's adjustment by producing external conflict. The concept of youth culture as developed by Talcott Parsons is presented at length, together with some reference to the writing of Kingley Davis. Case-studies illustrate certain of the principles deduced from social theory. There is unfortunately no reference to the relevant findings of the California growth-study of adolescents.

Psychologists may be most interested in the chapter on therapeutic management of anxiety in children, which is an attempt to bring the points of view of Dollard and

Miller and of Mowrer to bear on the problem of child-therapy. The chapter recognizes the problem of permissiveness in therapy; while permissiveness may be necessary early in therapy to enable the patient to free himself from morbid anxiety (to unlearn it), beyond that point permissiveness becomes disserviceable because it provides no cues or rewards for the new learning which must follow.

People well grounded in the literature of child-development will find in this text much that is familiar. They will regret that, as a compendium of the contribution of the behavioral sciences to psychoanalytically oriented child-guidance workers, it omits so much relevant material on developmental trends in behavior, the significance of social skills for the adjustment, and the cumulative effect of environmental stimulation on behavior. Perhaps we chiefly regret that the workers to whom the text is addressed apparently have such little exposure in training to child-development materials!

University of Minnesota

DALE B. HARRIS

Experimental Hypnosis. Edited by L. M. LECRON. New York, Macmillan, 1952. Pp. xviii, 483.

Every thirty years, according to the editor, there is a resurgence of interest in hypnosis. This book reflects the fact that we are at present in the midst of such a resurgence, which may be seen in the number of professional societies springing up, and in the number of books, journals, and articles concerned with research in the field of hypnosis. Whether or not this renewed interest will be temporary or permanent depends among other things on the contribution a book, such as this one, makes. This book deals with such topics in the field of hypnosis as: pre-Mesmer history; techniques of induction; tests predicting susceptibility; age-regression; relation to narcosis; influence on somatic function; antisocial uses; relation to extrasensory perception; dissociation; applications to obstetrics, dentistry, and psychotherapy; projective hypno-analysis; and time-distortion. Each chapter is introduced with a brief statement by the editor which indicates something of the context of the problem discussed. While the contributors writing for the book are mainly psychologists and psychiatrists from the United States, contributions from France, England, South Africa, and Denmark are included. References at the end of each chapter constitute a bibliography of much of the recent work in the field of hypnosis.

On the negative side one may object to such statements as that 69% of a group of 50 Ss remembered the day on which Christmas fell when they were regressed to the age of four. One is more than mildly surprised at this finding inasmuch as few four-year-olds have such information and since knowledge about the days of the week is an intelligence-test item scaled at the age of eight. Then again in discussion dealing with the same topic, one feels uneasy on encountering the statement concerning regression in Rorschach responses which may have high antecedent probability but little empirical validity. In another chapter, one reads that most coöperative individuals are quickly and easily hypnotized, while a few pages further on one is warned of the difficulty of hypnotizing "overactively" coöperative individuals. Reading the chapter on the use of hypnosis in dentistry, written by a psychologist, one wonders what rôle the novocaine played; ostensibly it was merely given to produce a feeling of "numbness." The discussion of the relationship of hypnosis and ESP makes little if any contribution. The reason for its inclusion, or so the reviewer believes, is that

the editor's "experience with ESP is nil." In general, many statements are made without supporting data and important experimental detail is often lacking, yet in some chapters there is undue emphasis on minutiae.

On the positive side are the many excellent chapters containing among other things practical assistance for those contemplating the professional use of hypnosis. A brief and concise history of hypnosis in the pre-Mesmer period is given. In another chapter the question of relationship between suggestibility and hypnosis is discussed; it is shown that suggestibility is not a unitary variable and that consequently the problem of the relationship has to be phrased so as to ask what kind of suggestibility is being referred to. In still another chapter a clear picture is given of the relationship between hypnosis and narcosis. Included in the book is a report of experiments, both successful and unsuccessful, encountered in the European literature (much of which is not yet available in English), of the effect of hypnosis on physiological functions such as chronaxie, leukocyte-count, and blood-sugar. There is also an interesting and detailed account of the use of hypnosis in obstetrics. This chapter, written by an obstetrician and a psychologist, is concerned with the use of hypnosis for the *average* female (reviewer's italics). The authors' thesis is that the fear of childbirth (engendered by our society) results in anxiety and tension and that these factors lower the limen of pain. Alleviation of the fear and tension involves the use of suggestion in whatever hypnotic state may be induced in the patient. On the basis of their findings, with experimental and control groups, the authors claim a decrease of about 20% in the time of labor and in the amount of drugs; they believe that their technique represents a logical extension of the Read-method of painless childbirth. A chapter written by a dentist practicing in a small town indicates the uses and limitations of hypnosis. It is of interest to note that as a result of the use of hypnosis, this dentist's clientele has grown and not, as often feared, decreased. The author of the chapter on antisocial acts in hypnosis is one who used to be "on the other side of the fence." He discusses the question of why both positive and negative results are reported and describes his own experimental work (which explains in part why his point of view changed).

One may say, in summary, that this book reflects the use of a scientific procedure which is more rigorous than formerly, that it represents an attempt to apply research findings to the average rather than to the exceptional individual, and, finally, that it does not hold up hypnosis as an experimental or therapeutic procedure without equal but rather as a useful technique to be used judiciously. This book may help to dispell the murky and ambiguous situation existing today in many of our colleges and universities where hypnotic research is considered 'unwise' if not completely prohibited. LeCron and the Macmillan Company are to be congratulated for publishing this book which demonstrates again that what hypnosis requires is not banning but investigation.

The State College of Washington

F. L. MARCUSE

Readings in Personnel Administration. Edited by PAUL PIGORS and CHARLES A. MYERS. New York, McGraw-Hill, 1952. Pp. xii, 483.

Teachers and students of industrial and business psychology will welcome this volume of readings by Pigors and Myers, not because of the 'psychology' it contains, but because it provides in readily accessible form the philosophy and principles

behind present-day human-relation practices in business and industry. Academic psychologists sometimes feel uneasy as they pursue methods of improving tests, merit ratings, training procedures, and the like, wondering if industry is really interested in the products of their work. This collection of some 45 different papers should ease that feeling and convince them that there is a tremendous need for and interest in their contributions. Indeed, psychologists *are* making contributions, for in this book on personnel administration one finds important achievements by A.P.A. members Douglas McGregor, N. R. F. Maier, Alex Bavelas, Mason Haire, R. N. McMurry, Guy W. Wadsworth, Jr., Alvin Zander, A. J. Cardall, Edward N. Hay, and Ordway Tead. By and large, the collection of articles is an excellent one, although we cannot agree with the statement on the dust cover that the volume is "essentially a collection of 'classics.'" It seems more accurate to say that the best of current thinking in various broad areas of personnel administration is collected—*current* because only two articles were written before 1940 and scarcely more than a half dozen before 1945. Over 20% of the articles were written in 1950.

The authors have avoided, in general, articles dealing with techniques in personnel administration, and a few papers more technical and empirical than the rest are made more refreshing by contrast. These include Maier's stimulating paper on the use of rôle-playing in supervisory training and McMurry's paper on the patterned interview as an instrument of selection. Perhaps the outstanding article in the collection, from the psychologist's point of view, is Mason Haire's discussion of psychological testing-programs for selection of employees. Haire's cautions in the use of tests should be understood by every student of industrial psychology and certainly by the management of every company which uses these devices. In selecting by means of tests of ability, Haire points out, "... an employer finds himself at a peculiar disadvantage because he no longer understands, in the same way as before, why he hires or rejects a given individual . . . It is one thing to say of a rejected applicant, 'Him? He couldn't even back up a truck!' and quite another to say, 'He scored lower than 38% of the applicants on a battery of tests that has been shown to have a .45 correlation with success on the job . . .'" (p. 316-317). Solomon Barkin's paper entitled *A Trade Unionist Appraises Management Personnel Philosophy* expresses a point of view which must be recognized and dealt with in any human-relations program. In a clear and concise way organized labor's objection to management's 'humanistic' approach is expressed.

In summary, those interested in the applications of psychology to business and industrial problems will find this collection of papers extremely helpful. They will find almost 50 expressions of contemporary views concerning the general principles and philosophy of business administration assembled for their convenience under six major headings: (1) the nature and scope of personnel administration; (2) analyzing and handling personnel problems; (3) the foreman: his place and his problems; (4) building and maintaining work teams; (5) wages and work assignments; (6) employee services and programs. While variations in quality exist among these papers, each is worthy. Designed as a supplementary text for courses in personnel administration, a function it fulfills exceedingly well, the collection will certainly find its way into lectures and outside reading assignments for students of business and industrial psychology.

Moosehaven Research Laboratory

ROBERT W. KLEEMEIER

Industrial Psychology. By JOSEPH TIFFIN. Third edition. New York, Prentice-Hall, 1952. Pp. 559.

The basic objective of this third edition of Tiffin's text is to summarize in some detail the various methods that have been, and are being, used in applying psychological techniques to problems encountered in the field of industrial psychology. The wide experience of the author as a consultant has no doubt directed the emphasis on techniques which have helped to gain for industrial psychology the recognition of industrial management. Giving adequate coverage to this area has left little space for descriptions of experimental design in industrial research. Some teachers of industrial psychology may deplore, too, the lack of attention given to the social psychological aspects of industrial behavior; the book devotes little space to such topics as group-dynamics, personnel-counseling, and leadership.

The general organization of *Industrial Psychology* remains essentially the same as in the second edition, with 5 of the 15 chapters given over to principles and descriptions of testing. A new introductory chapter has been added which describes briefly those developments which have influenced the acceptance of psychological methods by industry. Expansion of such topics as morale, interviewing, and supervision is an improvement over the earlier editions, but little has been added to the treatment of efficiency, accidents, and safety. The condensation of the chapter on visual skills is well done and might have set a pattern for condensing some of the material on testing, still leaving much of the clear and useful information on tests. Space might thus have been saved for the expansion of the sections on training and job-evaluation, and for the inclusion of recent findings in such important areas as conference-techniques. Some readers should find the appendixes on statistics, Taylor-Russell tables, and a selected list of test-publishers of value. A good job has been done on the selection and printing of graphs throughout the book, and these complement the text nicely.

In summary, this book has its greatest value as a text where the instructor wishes to lay emphasis on techniques. For a broader coverage of the field of industrial psychology in the classroom, the book will have to be supplemented. The industrial practitioner will find this edition most useful, as he has the earlier editions.

Carnegie Institute of Technology

B. VON HALLER GILMER

Principles of Human Relations. By NORMAN R. F. MAIER. New York, John Wiley and Sons, 1952. Pp. x, 474.

"This book is written for persons who are interested in human relations problems. . . . The examples and applications are pointed towards industry, but the principles apply to all situations where leaders must deal with groups or individuals." Maier's objective leads, inevitably, to a fluctuating style. The book reads, at times, as if it was intended for college students, at other times as if it was aimed at management, and at still other times as if it was a teacher's workbook. This fluctuating style stems, in part, from the fact that about 20% of the material has been reprinted, with slight changes, from industrial and human-relations journals.

Maier begins by pointing up industry's need for human-relations programs, and by indicating briefly what human relations is, what it can do, and how it is done. He then outlines his industrial training program in human relations—a program in which roughly 40% of the time is spent on such basic principles of psychology as

motivation and individual differences, 40% on techniques of democratic leadership and of non-directive counseling, and 20% on special problems such as frustration, morale, and fatigue. The next three chapters are devoted to methods of group-discussion and to rôle-playing techniques in small and large groups. After reporting some reactions of industry to human-relations training (mostly favorable), Maier specifies, in detail, the leader's procedures in reaching group-decisions, and adds two chapters of case studies for instructional use. Then there is a chapter on the leader's influence on the quality of group-decisions, illustrated with transcriptions of two problem-solving conferences in which the quality of the group-decision differed because of the leaders. Finally, Maier outlines techniques for the supervisor to use in his rôles of interviewer and counselor.

One of the two really outstanding characteristics of Maier's book is the systematic presentation, for the first time, of the techniques, objectives, and uses of discussion-methods, rôle-playing, and group-decision procedures. In most cases, these presentations are summarized in a 1-2-3 fashion. The second outstanding characteristic is the abundant use of illustrative case-studies, to which almost 45% of the book is devoted. The last chapter, too, is a very fine introduction to the techniques of non-directive counseling.

Some psychologists, who are familiar with Professor Maier only through his experiments in reasoning, will be interested to see, particularly in his chapter on "The Quality of Group Decisions as Influenced by the Discussion Leader," the continuity between Maier's earlier work and his present interests. Many teachers will find the educational implications of the systematic discussions profitable and interesting. Although the book is not particularly well adapted to low-level courses, it will be well worth considering for upper-level service courses and for adult courses.

Knox College

ROBERT S. HARPER

Personality Development. By J. S. SLOTKIN. New York, Harper, 1952. Pp. x, 390.

The author sets himself the gargantuan task of studying the effects of culture and society on the development of personality. In order to do this, he first develops "a systematic theory of personality" drawing on the fields of anthropology, biology, psychiatry, psychology, and sociology. Since the author is a social anthropologist, he relies mainly on social and cultural influences.

The book is divided into four sections headed: Inheritance, Socialization, Cultural-ization, and Individualization. Each of these sections includes an account of the developmental sequence of the first five years, since it is during this period, Professor Slotkin feels, that the basic personality pattern is established.

Unfortunately, after several readings of the book, one fails to find the "systematic theory of personality" promised on the fly leaf. Rather, due to the unusual organization of the book, one is soon bogged down in a morass of simple statements about personality or a given phase of development, each followed by a footnote giving the source of the idea. These references are often to psychological experiments. Each statement is then followed by a series of quotes varying in length from one line to several pages. The sources of these quotes run the gamut from the author's own notes about his personality and lengthy historical accounts of wolf-children to experimental data. In addition, each chapter is followed by a series of "Historical

Notes" running from a half page to twelve pages. The net result is that one is extremely impressed with Professor Slotkin's wide acquaintance with the literature, and literature in general, but can only feel that more time should have been spent thinking and less reading. The result of all this reference material seems to be only the restatement throughout the book that personality is the sum total of behavior patterns derived from learning and building on inheritance—hardly a systematic theory of personality development out of which "hitherto unexplored areas for new research" will open up.

Bryn Mawr College

DONALD R. BROWN

A Further Study of Visual Perception. By M. D. VERNON. New York, Cambridge University Press, 1952. Pp. xi, 289.

It is refreshing in these times to find a book on perception which, while not slighting certain currently popular problems, goes beyond the effect of needs and wishes, or pretty demonstrations of well known determinants of size and depth. In the introduction, Miss Vernon remarks on the changes in the field of perception which have taken place since the appearance of her *Visual Perception* in 1937, a book that was influenced largely by the Gestalt school. The present work was written from an essentially eclectic point of view, in the belief that "the apparently clear outlines originally drawn by the Gestalt psychologists have become blurred and dimmed." The principal defect of the new book is its brevity; a great number of researches are cited with little attention to details of method, data, or theoretical setting. It might be well, too, if more consideration had been given to the neurological substrate of perception. Nevertheless, the book should prove useful both as a text and as a reference work.

University of Texas

JOHN KRAUSKOPF

In the Name of Science. By MARTIN GARDNER. New York, G. P. Putnam's Sons, 1952. Pp. x, 320.

This is a remarkably good book for the public, which confuses cult with science, for the teacher, who wishes to give his students interesting examples of unscientific method, and for the social psychologist, who is concerned with the formation of in-group prejudice. The author writes clearly and with scientific validity from a great fund of odd formation. He discusses Voliva's flat earth, Teed's hollow earth with all of us on the inside, Velikovsky's worlds in collision, Charles Fort's Fortean Society, the flying saucers, the new mechanics and cosmology of A. W. Lawson's lawsonomy, Roger Babson's Gravity Research Foundation which seeks to discover a gravity screen, dowsing (with especial attention to Kenneth Roberts' Henry Gross), George McCready Price's new geology, inheritance of acquired characters from Lamarck to Lysenko, race supremacy with reference of many protagonists including Hitler, Donnelly's theory of Atlantis, the convictions of John Taylor and C. P. Smyth about the occult significances of the Great Pyramid, Dr. Bates on perfect sight without glasses, medical cults and quacks and food faddists too numerous to specify, Wilhelm Reich's orgonomy, Hubbard's dianetics, the non-Aristotlean therapy of Korzybski's general semantics, phrenology from Gall on, graphology, and finally, with a good word for its sincerity and a doubt about its cultish status, Rhine's ESP and PK.

Harvard University

EDWIN G. BORING

THE AMERICAN JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXVII

JUNE, 1954

No. 2

THRESHOLD FOR VISUAL FORM: A DIFFUSION MODEL

By M. E. BITTERMAN, JOHN KRAUSKOPF, University of Texas, and
J. E. HOCHBERG, Cornell University

The experiments here reported are concerned with the visual identification of geometric form. Much of the previous work on the problem has been purely empirical in character; many figures have been studied under a variety of conditions, but theoretical analysis has rarely been carried beyond the complacent rejection of certain preliminary notions about 'goodness' of form taken from the early Gestalt literature.¹ In many instances, no evidence for the differential discriminability of simple geometric forms has been found. For example, Helson and Fehrer,² in one of the more significant papers on the problem, reported different forms to be favored by different criteria and variations among Os so large as to make any generalization about 'goodness' impossible. In other instances, such as the more recent experiment by Casperson,³ significant differences in discriminability have been reported with little or no attempt to understand them in terms of underlying visual processes. In consequence, the choice of particular experimental methods and conditions often appears to have been quite arbitrary. Casperson has declared that "relative discriminability . . . is essentially unaffected" by the technique of measurement employed,⁴ but there is little evidence for this belief and no logical justifi-

* Accepted for publication June 2, 1953. The research here reported was conducted at the University of Texas under terms of USAF Contract No. AF 33(616)-63. The project was administered by the Psychology Branch, Aero Medical Laboratory, Directorate of Research, Wright Air Development Center, with Dr. Kenneth T. Brown serving as Project Engineer.

¹ R. B. Sleight and Grace Duvoisin, *An Annotated Bibliography of Form Perception*, Office of Naval Research, 1952, 1-28.

² Harry Helson and E. V. Fehrer, The rôle of form in perception, this JOURNAL, 44, 1932, 79-102.

³ R. C. Casperson, The visual discrimination of geometric forms, *J. Exper. Psychol.*, 40, 1950, 668-681.

⁴ *Ibid.*, 668.

cation. It is not clear, for example, why the campimetric method⁵ should be expected to yield results comparable to those obtained with foveal presentation at low levels of illumination.⁶ The present experiments, in any event, were neither atheoretical nor methodologically arbitrary. They were derived from a theory which suggested not only the general need to reinvestigate the discrimination of form but the method to be employed as well.

Central to the Köhler-Wallach explanation of figural after-effects is the postulate that a visual stimulus produces an alteration in the chemical condition of the tissue fluid which surrounds the cells of the visual cortex.⁷ It is as a function of differences in potential between regions of different activity (ionic concentration) that the figural currents arise which are assumed to determine perceptual experience. Thus, if one area of the retina, f , is the site of more intense stimulation than the surrounding area, g , a greater number of nerve impulses are conducted along the pathways originating at f than along those originating at g . In the course of their activity, the nerve fibers are assumed to liberate chemical substances at their termini, the rate of secretion being proportional to the frequency of impulses. If F is the region of the visual cortex at which terminate the pathways arising from f , and if G is the cortical region corresponding to g , more intense stimulation of f than of g will result in a greater frequency of impulses at F than at G . Corresponding to the activity at F and G , two concentrations of chemical substances, F' and G' , will develop. If these substances dissociate into ions, a potential-difference will be created between F' and G' in relation to which a current will flow through the cortical tissue (assumed to function as a volume conductor).

It is reasonable to suppose that the difference between the concentrations F' and G' must exceed a given value before the difference in potential between the two regions is sufficient to produce a flow of current which will result in figural perception. If stimulation is confined to the retinal region f , as in the case of a luminous figure in a dark room, this threshold difference in concentration can be expressed in terms of the magnitude of F' alone. The F' concentration depends, however, not only on the intensity of stimulation at f , but also on the effectiveness of coëxisting dissipating processes such as diffusion and enzymatic action. When the process of diffusion is considered, without the assignment of specific values to the variables involved, certain implications for the perception of form become evident.⁸

⁵ R. M. Collier, An experimental study of form perception in indirect vision, *J. Comp. Psychol.*, 11, 1931, 281-289; Nathaniel Kleitman and Z. A. Blier, Color and form discrimination in the periphery of the retina, *Amer. J. Physiol.*, 85, 1928, 178-190; N. L. Munn and G. A. Geil, A note on peripheral form discrimination, *J. Gen. Psychol.*, 5, 1931, 78-88; C. A. Whitmer, Peripheral form discrimination under dark adaptation, *J. Gen. Psychol.*, 9, 1933, 405-419; M. J. Zigler, Barbara Cook, Dorothy Miller, and Luisita Wemple, The perception of form in peripheral vision, this JOURNAL, 42, 1930, 246-259.

⁶ Helson and Fehrer, *op. cit.*, 79-102.

⁷ Wolfgang Köhler and Hans Wallach, Figural after-effects, *Proc. Amer. Phil. Soc.*, 88, 1944, 269-357.

⁸ This discussion of cortical events reads (for the sake of simplicity) as though

Since the perceived figure is assumed to be isomorphic with the current flow around F' , and since F is assumed to be isomorphic with f , the stimulating figure will be correctly perceived only if F' is isomorphic with F . It may be expected that $F'o$, the original configuration of F' , will be determined by the configuration of F . If, however, $F'o$ is changed by diffusion to $F'd$, the correspondence with F will be lost and the accuracy of perception will thereby be reduced. Perceptual error should not, of course, be random. On the assumption that it results from the discrepancy between $F'o$ and $F'd$, error should be predictable from the process of diffusion which is responsible for the discrepancy.

Since the loss occurs largely at the boundaries of $F'o$, the amount and pattern of diffusion will vary with the extent and configuration of those boundaries. The amount of diffusion should be closely related to the perimeter of $F'o$ —the greater the perimeter, the more the diffusion. The pattern of diffusion (as distinct from the amount) should depend upon the configuration of $F'o$, and two dissimilar figures of equal perimeter and area should therefore be affected quite differently by diffusion. In general, corners should be rounded, gaps should be closed, asymmetry and inhomogeneity should be reduced, with small details tending to disappear more rapidly than larger details. It also may be assumed that an increase in the frequency of impulses arriving at F will counteract the effects of diffusion and thereby improve the accuracy of perception. Given a medium of constant viscosity and some limit on the rate of enzymatic destruction, an increase in the frequency of impulses at F should bring the configuration of F' into closer correspondence with the configuration of F .

These considerations lead to some rather specific predictions for a simple experimental situation—the measurement of form thresholds in terms of intensity of illumination required for the perception of luminous figures briefly exposed in a dark room. Thus, the threshold of the circle should be lower than that of any other form of equal area. Loss by diffusion should be least in the circle since it has the smallest perimeter; furthermore, whatever diffusion does occur should leave the form of $F'o$ unaltered, since the pattern of diffusion from a circular region is circular. When area is increased with form held constant, threshold may be expected to decrease, since the relative amount of diffusion will be less for large figures than for small (the ratio of perimeter to area varies inversely with area), and since the magnitude of critical detail increases with area. The theoretical considerations here described led Hochberg, Gleitman, and Macbride to seek (and to find) evidence of a positive correlation between form threshold and a ratio of perimeter to area

F' and G' were assumed to be homogeneous and as if the boundary between them were characterized by a sudden increase in density. Actually, of course, neither of these assumptions is made. As the concept of diffusion itself suggests, a contour must be represented by a gradient of density.

(P/A).⁹ It should be noted that such a relationship is predicted only when magnitude of critical detail is controlled. In certain groups of figures, the relationship between form threshold and P/A will almost certainly be inverse. For example, in a series of regular polygons equal in area and differing in number of sides (such as square, pentagon, hexagon, . . . decagon)—where variation in P/A is small as compared with variation in the size of critical detail— P/A decreases as the number of sides increases, but form threshold may be expected to increase. The theory suggests that the problem of critical detail may itself be reduced to an effect of P/A for some restricted region of the figure, but no reasonable way of segregating a particular part of a figure for purposes of computation has been found. Further consideration of this problem might best be postponed until the broader features of the model have been evaluated.

Examination of the literature reveals few studies bearing upon the validity of these deductions which, it should be emphasized, apply to specialized experimental conditions. One of the most important studies is that of Helson and Fehrer who measured foveal thresholds (in terms of intensity of illumination) for six figures of equal area: triangle, rectangle, square, circle, semicircle, and angle (a square from which a smaller square had been cut away).¹⁰ Three thresholds were distinguished—the intensities required for "just perceived light," "just perceived form," and "certain form." The results of the experiment did not seem to warrant generalizations about 'goodness' of form (except perhaps the generalization that the circle was no 'better' than any other figure), and the authors accused the configurationists of having "over-emphasized and over-generalized the importance of form as such" (p. 102). Helson and Fehrer noted the fact that light was almost always seen before any trace of form was visible—that "nearly twice as much light" was necessary "to perceive forms in a hazy, indistinct manner" as to "barely perceive light"—and concluded: "The assertion that forms are first and primary in perception therefore falls to the ground" (p. 101).

The first question to be asked with respect to these results is why a consistent relation between form threshold and perimeter did not appear, and the answer is given in methodological terms. Since determinations were made by the method of adjustment, real differences among the forms may have been obscured by random errors stemming from the failure to control time of exposure. The diffusion model suggests that it would have been better in such an experiment to use brief exposures at successively higher levels of illumination instead of continuous variation. A much more important criticism, however, centers on the fact that repeated determinations were made, in consequence of which the *O*s may have been influenced unduly by their knowledge of the range of forms employed. As Drury noted in a discussion of the early campimetric experiments,¹¹ "It is difficult, if not

⁹ J. E. Hochberg, H. Gleitman, and P. D. Macbride, Visual thresholds as a function of simplicity of form, *Amer. Psychol.*, 3, 1948, 341-342.

¹⁰ Helson and Fehrer, *op. cit.*, 79-102.

¹¹ M. B. Drury, Progressive changes in non-foveal perception of line patterns, this JOURNAL, 45, 1933, 628-646.

impossible, to estimate the effects of recognizing, instead of 'seeing,' the known forms. To touch off a verbal name or other epithet, it is only necessary that the stimulus-pattern should be sufficient to suggest 'that form.' Practice and repetition make *O* more and more expert in distinguishing the several members of the series (seen at times with full foveal adequacy). The experiment comes to be, therefore, more a process of coaching in the interpretation of indistinct signs than a discovery of retinal adequacy at various removes from the fovea." Drury's criticism was directed mainly at the use of conventional forms; but it may be argued that the *bête noir* of such experiments is not as much pre-experimental familiarity with the stimuli as it is the knowledge that the stimulus encountered on any trial can have been drawn only from a small, delimited group.

Evidence that the functional task of the *O*s of Helson and Fehrer was simply to determine which of six possible forms was presented on a given trial is to be found in a table (p. 94) showing the number of times each figure was mistaken for every other in the first 50 determinations of the threshold for "just perceived form." *Not a single figure* was ever in these trials designated as *anything* but one of the six figures employed in the experiment. Helson and Fehrer themselves recognized the distorting effect of knowledge when they decided to consider only the results of ascending determinations—"knowledge of the forms prevented the *O*s from reporting when they actually disappeared in the descending series" (p. 85). The belief that knowledge produces distortion also in ascending series is supported by a comparison of the results of Helson and Fehrer with those of Hochberg, Gleitman, and Macbride.¹² The latter investigators, in a relatively crude, group experiment, computed thresholds for three conventional forms (circle, square, and cross of equal area) on the basis of *one* ascending determination per figure for each *O*—a procedure which ruled out the influence of knowledge of the forms. Significant increases in threshold with perimeter were found.

The second question to be asked about the results of Helson and Fehrer concerns the lack of a consistent relationship between perimeter and threshold for brightness ("just perceived light"). Whether or not such a relationship is deduced from the diffusion model depends on the rate of diffusion that is assumed. If the rate of diffusion were negligible, neither form nor brightness thresholds would be expected to vary with perimeter. With an intermediate rate of diffusion, a relation to perimeter would be expected for form thresholds but not for brightness thresholds. Since the loss occurs mainly at the boundaries of a figure, small variations in the amount of diffusion might be critical for the integrity of the boundaries but without effect on the concentration at the center of the figure (which probably determines whether the brightness threshold will be reached). With still higher rates of diffusion, both form and brightness thresholds would be expected to increase with perimeter. The negative results of Helson and Fehrer are, therefore, not crucial for the diffusion model at its present level of development. It might be well, however, to repeat the experiment under conditions in which time of exposure is better controlled.

A third question posed by the results of Helson and Fehrer concerns the perception of form at the brightness threshold. Wilcox attributed the finding that light was perceived before form to a confusion of objective and experienced form; he

¹² Hochberg, Gleitman, and Macbride, *op. cit.*, 341-342.

denied the existence of unformed perceptual fields and insisted that his own *O*s "invariably" described as circular a variety of figures presented at the brightness threshold.¹³ This difference of opinion centered about two issues which a subsequent discussion by Helson did not fully resolve.¹⁴ First there is the question of whether a patch of light seen against a dark ground can be regarded as unformed at the brightness threshold; this question is perhaps not entirely terminological, but it is largely irrelevant to the present investigation. Then there is the question of whether the patch of light inevitably looks circular as soon as it can be regarded as formed; the diffusion model suggests a negative answer to this question. It is true that the discrepancy between *F'o* and *F'd* is smallest for circular stimuli, and it is true that *certain* forms (e.g. regular polygons) should tend to be perceived as circular before the threshold for correct identification is reached. The diffusion model does not, however, lead to the conclusion that *F'd* must inevitably be circular at the brightness threshold irrespective of the rate of diffusion and the configuration of *F'o*.

Another experiment which has a direct bearing on the validity of the diffusion model is Fry's exploration of configurational influences on the discriminability of differences of brightness between adjacent areas.¹⁵ The method was to project a figure on an illuminated ground and to measure the intensity of the figure required to differentiate it from the ground. A number of Fry's findings have interesting implications for the diffusion model. For example, a wavy or saw-tooth border having a fine pitch was seen as a continuous straight edge at threshold level, and a regular polygon (such as a hexagon) appeared as a disk at threshold level; in each case, it should be noted, the segments were several times larger than the diameter of the blur circle. These results clearly suggest the operation of a diffusion process.

EXPERIMENT

In the experiment to be reported, foveal form thresholds for a variety of figures differing in form and area were determined. Luminous figures presented in a dark room were exposed at successively higher levels of intensity until form was correctly perceived. The experiment was patterned after that of Helson and Fehrer from which it differed in two important respects. A succession of brief exposures of the stimuli replaced the continuous method of the earlier study, and the *O*s were ignorant of the

¹³ W. W. Wilcox, Helson and Fehrer on The rôle of form in perception, this JOURNAL, 44, 1932, 578-580. The extent to which the *O*s of Helson and Fehrer fell into the stimulus-error (as Wilcox charged) can be estimated from data on the frequency with which forms other than those employed in the experiment were named as the "just perceived level" (Helson and Fehrer, *op. cit.*, 94). The lack of such reports suggests that the *O*s were more concerned with estimating the nature of the stimulus than with describing their experience.

¹⁴ Helson, Dr. Wilcox on The rôle of form in perception, this JOURNAL, 45, 1933, 171-173.

¹⁵ G. A. Fry, The relation of the configuration of a brightness contrast border to its visibility, *J. Opt. Soc. Amer.*, 37, 1947, 166-175.

forms employed. (As in the study by Hochberg, Gleitman and Macbride, only one determination for each form was made with each *O*). The purpose of the experiment was to determine whether meaningful differences in form threshold would appear under these conditions.

Observers. Twenty-four young men and women, students at the University of Texas, served in the experiment. They were psychologically unsophisticated, and their vision, by their own reports, was either normal or corrected to normal.

Apparatus. A modified Nagel adaptometer was employed for the experiment (Fig. 1). Illumination was provided by three 25 w. tungsten lamps operated at 115 v. by means of a voltage-regulator. The adaptometer was equipped with three neutral density filters, each of which transmitted 5% of the incident light. To eliminate rod vision, a red filter was employed; its transmission was 5.2% and it

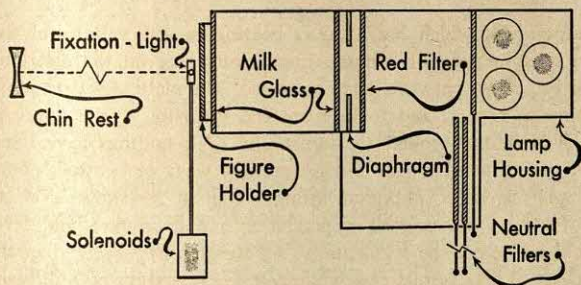


FIG. 1. SCHEMATIC DIAGRAM OF THE APPARATUS

plotted on the spectral locus at $643\text{m}\mu$. A linear diaphragm was used to which a scale measuring percentage of maximal illumination was coupled, and a small lamp was housed in the apparatus to permit the scale to be read by *E* in the dark experimental room. A figure-holder, into which the stimulus-figures could be inserted individually, was fixed to the front face of the apparatus. The instrument was calibrated with a Luckiesh-Taylor meter. Readings were made without the red filter and then corrected for its transmission. With the diaphragm opened maximally (100 on the linear scale and 2 on the log scale), the luminance was 0.004 ft-L.^{16}

A chin-rest secured the head of *O* who sat at a distance of 5 ft. from the stimulus-figure which was viewed binocularly. A tiny neon lamp shining through a pin-hole in a small box, provided a reddish fixation-point. The box was attached to the end of a metal rod in such a way that the fixation-point was situated directly in front of the center of the figure-holder. The rod was mounted on a ball-bearing and could be rotated out of the line of sight by the action of two solenoids. The sequence of events on each trial was controlled by a bank of three interval-timers

¹⁶ We are indebted to Dr. Harry Helson for advice on the calibration of the apparatus.

of the Hunter-Brown type.¹⁷ After a ready-signal, *E* threw a lever which extinguished the fixation-light and moved it out of the field of view. After an interval of 0.5 sec., the adaptometer-lamps were turned on to expose the stimulus-figure for a period of 0.5 sec. Then an interval of 2 sec. elapsed, following which the fixation-light was brought back into position and turned on once more.

Stimuli. The figures used in the experiment were cut from heavy paper. (Ma-

TABLE I
THE FIGURES EMPLOYED

Form	Area (Square inches)	Perimeter (Inches)	Perimeter/Area
Circle	.25	1.77	7.08
Square; diamond*	.25	2.00	8.00
Square; diamond*	1.00	4.00	4.00
Triangle	.25	2.27	9.08
<i>L</i>	.25	2.31	9.25
Cross; <i>X</i> *	.25	2.69	10.75
Cross; <i>X</i> *	1.00	5.36	5.36
<i>T</i>	.25	2.69	10.75
<i>H</i>	.25	3.02	12.10

* Identical forms differing in orientation by 45°.

chined metal figures were prepared for preliminary explorations before it was found that sufficient accuracy could be obtained from paper.) With these templates, contact-negatives were made—transparent figures on opaque grounds. The negatives were bound to thin plates of clear plastic which could be inserted individually into the figure-holder on the adaptometer.

The figures employed are described in Table I. There were nine different forms (circle, square, diamond, equilateral triangle, *L*, cross, *X*, *T*, *H*) of which all but diamond and *X* are illustrated in Fig. 2. The diamond was a 45° rotation of



FIG. 2. THE FORMS USED IN THE EXPERIMENT

the otherwise identical square, and the same relation obtained between cross and *X*. The *T* was constructed to have the same perimeter as the cross and *X* of equal area. The *L* was not a conventional one, since it was designed to have a perimeter similar to that of the triangle. Each form was constructed with an area of 0.25 sq. in.—the *small* figures. In addition a diamond and an *X* equal in area to 1.00 sq. in. (*large* figures) were prepared. The figure with the greatest lateral extent was the large *X* which measured 1.34 in. For the *O* seated 5 ft. away, its angular size was 1°16'.

Procedure. Each *O* was brought into the air-conditioned dark room employed for experiment, dark-adapted for a period of 10 min., and given essentially the following instructions:

¹⁷ T. A. Hunter and J. S. Brown, A decade-type electronic interval-timer, this JOURNAL, 62, 1949, 570-575.

This is an experiment in visual form perception. You are going to be shown some simple figures at very low brightnesses, and your task will be to identify each figure correctly. Fixate the reddish point of light when I say 'ready,' and when you can see it clearly, answer 'ready.' When you give your ready-signal, I shall turn out the fixation-light, and a half-a-second later a dimly illuminated figure will appear where the fixation-light was. The figure will stay on for half a second, and you are to describe its appearance. On the first trial you will not see anything, but on succeeding trials I shall increase the brightness of the figure until you can name it correctly. We will then repeat the procedure for several other figures.

For each *O*, one determination of the threshold for each figure was made in the following manner: The adaptometer scale was set below the brightness threshold for the first exposure. On successive trials the illumination was increased by steps of 0.05 log unit, and the series terminated when *O* correctly named or described the form. Most *O*s had no difficulty in naming the familiar forms employed, but some of the *O*s were troubled by one or another of the forms. For example, the diamond on occasion was not called by name on the assumption that one diagonal of a diamond should be longer than the other; the squat *L*, selected for a minimum *P/A* ratio, posed a naming problem for some *O*s; and even the cross (which should have been quite familiar by virtue of form and color—'red cross') caused a certain amount of difficulty. At times it was necessary for *E* to question *O* for the purpose of determining what was seen. If the cross was called a 'four-armed figure,' *E* asked about the relation of the arms, the orientation of the figure, and so forth. The criterion of threshold was, therefore, not as objective as might be desired, although difficulty of naming was encountered only in a minority of instances. Names for the forms could not be supplied to the *O*s in advance of the presentations because it was desired that they be ignorant of the forms to be seen. It was for the same reason that only one determination for each form was made with each *O*—an unusual and otherwise inefficient procedure. The importance of these precautions is demonstrated by the results of a later experiment in which repeated determinations were made.

Each *O* worked with nine of the figures. Thresholds for small circle, square, triangle, *L*, cross, and *H* were determined first in an order which varied systematically from *O* to *O*—a 6 x 6 Latin square was constructed and four *O*s were randomly assigned to each row. Next the threshold for the (small) *T* was determined for each *O*. Finally, six of the *O*s (selected at random) were presented with small *X* and large diamond, six with large diamond and small *X*, six with large *X* and small diamond, and six with small diamond and large *X*. Thus, for the group as a whole there were 24 determinations for each of the first seven figures (small circle, square, triangle, *L*, cross, *H* and *T*) and 12 determinations for each of the last four figures (small diamond, large diamond, small *X*, and large *X*). The last two determinations for each *O* provided (over the entire group) a self-contained study of the effect of area which was linked to the earlier series of figures in that small diamond and *X* were 45° rotations of the otherwise identical small square and cross.

Results. The outcome of an analysis of variance for the first six figures, which were incorporated in the Latin square, is summarized in Table II. Variance due to order of presentation was insignificant, but variance due

to trials reached significance at the 5-% level of confidence. Variance due to form (for these six figures of equal area) was highly significant.

In Fig. 3, mean thresholds for all of the figures used in the experiment

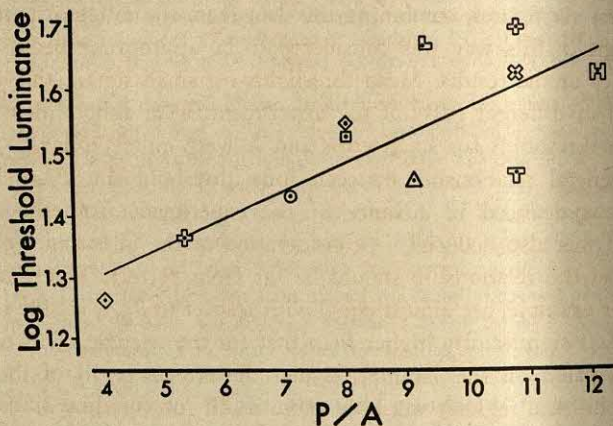


FIG. 3. FORM THRESHOLD AS A FUNCTION OF P/A

The two lowest P/A values represent the large figures and the others represent the small figures. The line was fitted to the data by the method of least squares.

TABLE II

ANALYSIS OF VARIANCE FOR SIX FORMS OF EQUAL AREA (CIRCLE, SQUARE, TRIANGLE, L, CROSS, AND H)

Source of Variation	Sum of Squares	df	Mean Square	F
Independent observations:				
Order of presentation	.3522	5	.07044	0.72
Residual between Ss	1.7588	18	.09771	
Total between Ss	2.1110	23		
Correlated observations:				
Form	1.5518	5	.31036	23.91*
Trials	.1598	5	.03196	2.46†
Residual from Latin square	.2857	20	.01429	1.13
Residual within Ss	1.1417	90	.01269	
Pooled error	1.4272	110	.01298	
Total within Ss	3.1390	120		
Total for experiment	5.2500	143		

* Significant beyond the 1-% level of confidence.

† Significant beyond the 5-% level of confidence.

are plotted as a function of P/A . A line of best fit, computed by the method of least squares, also is plotted. Inspection of the data reveals a

rather marked linear relationship between the two variables. The degree of relationship, expressed in terms of a rank-difference correlation between mean threshold and P/A is 0.78 (significant beyond the 1-% level of confidence). Despite the significant trials variance found in the analysis of the first six figures, combining the data from the different parts of the experiment in this way was considered to be appropriate because of the consistency of the results. Mean thresholds for small square and diamond, measured in different parts of the experiment, were almost identical, and the mean thresholds for small cross and X were quite similar.

The general relationship between form threshold and P/A shown in Fig. 3 was deduced in advance of the experiment from the diffusion model. It was also deduced, in consideration of the importance of critical detail, that the relationship should be far from perfect. The triangle and the L , for example, are almost equal with respect to P/A , yet the threshold for the L is significantly higher than that for the triangle. This difference was anticipated on the assumption that the critical detail of the L (the upper right-hand region) would be more readily obscured by diffusion than would the angular relations among the sides of the triangle. The critical detail of the L is about equal to that of the cross and X , and the figures have about equal threshold despite the discrepancy in P/A . The lower threshold for T than for cross and X (which are equal to T in P/A) also is understandable in terms of the size of critical detail, and the discrepancy between the triangle and the square can be explained in the same way. Unfortunately, however, no completely satisfactory method for the quantitative evaluation of the contribution of critical detail to form threshold has yet been found.

Important support for the diffusion concept is provided by the qualitative data of the experiment. Although systematic records were not kept in the beginning, the O s were found to be remarkably consistent in some of their pre-threshold descriptions of the various forms. The square was frequently called a circle at pre-threshold levels, and the triangle was also mistaken for a circle (although not as often as was the square). The cross was frequently called a diamond, the X a square, the T a triangle with apex down, the L a semicircle or half-moon, and the H (on occasion) a butterfly; the circle was rarely called anything but a circle. From these reports it may be inferred that, in general, as the diffusion model suggests, corners tend to round and gaps to close. These impressions are confirmed by the observations of trained O s. For example, the cross is quite frequently seen as a diamond at low levels of illumination; as the

level of illumination is increased, the sides of the diamond become more and more invaginated, and finally the cross emerges.¹⁸

In a recent experiment on the discrimination of form, Casperson ignored the perceptual significance of erroneous reports for which he introduced a statistical 'correction' on the assumption that they were artifacts of 'response probability.'¹⁹ The data of the present experiment do not warrant such treatment. The incorrect responses certainly did not appear to be arbitrarily emitted; instead, particular incorrect responses tended to be uniquely associated with particular stimulus-figures. Furthermore, the diffusion concept points to a meaningful relationship between specific errors and the stimulus-figures which called them out. Evidence for the perceptual significance of pre-threshold reports also is provided in the paper of Helson and Fehrer.²⁰ For example, in 200 determinations with the *square* there were 20 erroneous responses, and the incorrect response in each case was *circle*; in the same number of trials with the *angle* there were 80 errors of which 57 were *triangle* and 12 were *semicircle*. Neither the quantitative nor the qualitative aspects of these results seem predictable from the concept of 'response probability.'

A physical diffusion model. In the interpretation of the results of Experiment I, a number of assumptions were made concerning the configurational changes which might be expected from diffusion. A physical model was developed to evaluate these assumptions. After some preliminary trials with dyes, a bacteriological technique was adopted.²¹

A culture of *Serratia marcesens* is suspended in a nutrient agar medium. The suspension is poured into a Petri dish and allowed to solidify. The form to be investigated (e.g. a cross) is cut from filter-paper, dipped into a 10% solution of HgCl_2 , and then dried. Finally, the filter-paper figure is placed on the surface of the agar and the dish is incubated. The growth of the bacteria is inhibited by HgCl_2 . During the period of incubation the paper figure absorbs water from the agar and the inhibitor diffuses through the medium. The diffusion of the inhibitor is visible in terms of the density of bacterial growth (the relationship, of course, being inverse).

The results obtained with this technique are illustrated in Fig. 4. The paper figures appear black in the photographs since the Petri-dishes were interposed between the source of illumination and the camera. Zones of

¹⁸ It should be noted that there are differences among *O*s with respect to pre-threshold reports. Thus, some *O*s fail to see the cross as a diamond or the *X* as a square, although all report some 'filling-in' of the regions between the arms.

¹⁹ Casperson, *op. cit.*, 668-681.

²⁰ Helson and Fehrer, *op. cit.*, 94.

²¹ We are indebted for advice on the development of the model to Dr. Orville Wyss of the Department of Bacteriology of the University of Texas.

inhibition appear as lighter areas surrounding each figure, and regions of normal growth may be seen in the periphery. The samples of Fig. 4 illustrate a number of important principles. In general, corners tend to round and gaps tend to close. Under comparable conditions of diffusion, small details are more readily obscured than are large details, and the

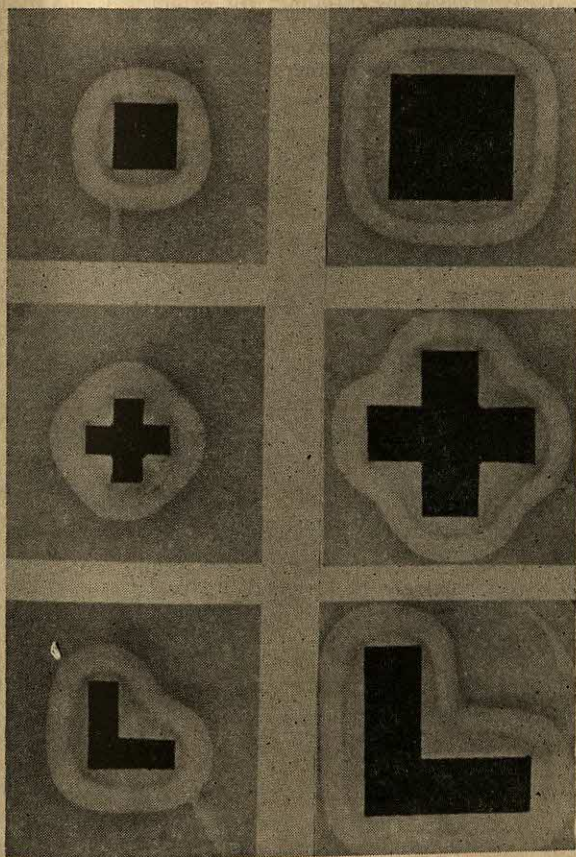


FIG. 4. RESULTS OBTAINED WITH THE PHYSICAL DIFFUSION MODEL

movement towards circularity is more apparent for some configurations than for others. Local effects are highly significant, as can be seen from a comparison of the small *L* (not the same *L* as employed in the experiment) and the large cross; the *L* is identical with the upper right quadrant of the cross. It should also be noted that in the case of the square and cross (two of the forms employed in the first experiment) evidence

for the validity of pre-threshold reports and of their interpretation in terms of a diffusion process is provided. The pattern of diffusion from the square approaches circularity, and the pattern of diffusion from the cross is at some point diamond-shaped. The diamond with invaginated sides produced by the large cross characteristically appears at levels of illumination intermediate between the threshold for brightness and the threshold for form.

The Osgood-Heyer model. The correspondence between the phenomenological data of the experiment and the qualitative evidence derived from the physical model does not, of course, demonstrate the existence of a common underlying process. It is, in fact, quite possible to interpret the psychological findings in terms of a conception very different from that which stimulated the experiment. In a recent paper, Osgood and Heyer²² have extended the statistical analysis of visual resolution by Marshall and Talbot²³ to provide an interpretation of figural after-effects which does less violence than that of Köhler and Wallach to the conventional view of cortical functioning, and the present results may be approached in much the same manner. It is possible, for example, that both the qualitative and the quantitative differences between square and cross which emerge in threshold determinations can be understood on the assumption that a contour is represented by a normal distribution of excitation which becomes increasingly peaked as the difference in intensity at the retina increases. There are, however, certain difficulties which must be met before this interpretation can be accepted. First, the spatial effects found in the experiment (e.g. the 'filling-in' of the large cross to produce a diamond-shaped figure) seem to be of a larger order of magnitude than those with which the Marshall-Talbot model was designed to deal. Second, if 'on-off' fibers are chiefly responsible for the distribution of excitation which represents a contour—an assumption required by the data of figural after-effects—and if phenomenal contour is determined by the maximum of the distribution, the systematic changes in shape which are found at low brightnesses do not seem readily predictable. It may be well, nevertheless, to explore in detail the implications of the statistical theory which maintains a concrete but flexible linkage with traditional neurophysiology.

²² C. E. Osgood and A. W. Heyer, Jr., A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1952, 98-118.

²³ W. H. Marshall and S. A. Talbot, Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity, in H. Klüver (Ed.), *Visual Mechanisms*, Biological Symposia VII, 1942, 117-164.

SUMMARY

Foveal form thresholds for luminous figures briefly exposed in a dark room were measured in terms of the intensity of illumination required for correct identification. In accordance with a diffusion model derived from the Köhler-Wallach theory of figural after-effects, form threshold was found to vary directly with the ratio of perimeter to area and inversely with the magnitude of critical detail. Data on the pre-threshold appearance of selected forms were related to results obtained with a physical diffusion model.

PERFORMANCE ON COMPLEX TACTUAL TASKS AFTER BRAIN INJURY IN MAN: ANALYSES BY LOCUS OF LESION

By JOSEPHINE SEMMES, SIDNEY WEINSTEIN, LILA GHENT,
and HANS-LUKAS TEUBER, New York University

In man and other primates, vision is commonly regarded as the dominant sense. It is therefore not surprising that laws of perceptual organization have been derived almost exclusively from studies of visual perception, and that investigations of memory and of 'higher' processes (problem-solving) have nearly always relied on visual presentation of material. Implicit in this preoccupation with vision is the assumption that findings in this modality can be applied to perception in general, independent of modality, and that students of problem-solving may safely disregard the visual nature of their tasks. Whether such generalizations from visual experience are justified is a question for empirical study rather than for *a priori* decision.

The problem becomes particularly important in studies of performance after cerebral lesions. In this area also, reliance has been primarily on visual tests for the demonstration both of perceptual deficits and disturbances of thought. Investigations in these fields have led to conflicting interpretations. Clinical reports describe visual, auditory, or tactile 'agnosias,' defined as a selective breakdown of recognition in a single modality (e.g. vision), in the absence of any significant sensory deficit for the modality in question. It has been frequently pointed out, however, that the evidence presented is not adequate to establish that the disturbance was as selective as claimed.¹ The opposite tendency appears in studies of intellectual functions after brain injury, in which altered performance usually is attributed to impairment of some general capacity such as 'abstraction' or 'categorizing.'² The tasks employed to demonstrate such a

* Accepted for early publication January 20, 1954. From the Psychophysiological Laboratory, Bellevue Medical Center. This research was supported in part by the Commonwealth Fund of New York and by the Office of the Surgeon General (Department of the Army), Contract DA-49-007-MD-312. Josephine Semmes served as post-doctoral fellow, U. S. Public Health Service, from 1952-1954.

¹ M. B. Bender and H. L. Teuber, Psychopathology of vision, *Progress in Neurology and Psychiatry*, 1949, 163-172; Josephine Semmes, Agnosia in animal and man, *Psychol. Rev.*, 60, 1953, 140-147.

² Kurt Goldstein and Martin Scheerer, Abstract and concrete behavior: An experimental study with special tests, *Psychol. Monog.*, 53, 1941, (No. 239), 151.

global change in intellectual function rely heavily on visual and verbal test-materials. Unless there is obvious aphasia or a gross defect of primary visual function in the *S*, the validity of these tests of such hypothetical capacities goes unchallenged. Subtle defects in these functions, which may interfere with *S*'s perception of, or reaction to the problem, may be insufficiently appreciated.

Studies of performance on complex tasks in more than one modality, and in patients whose primary sensory status has been explored in detail by quantitative rather than clinical methods, are needed to resolve the issues. Our current studies on effects of penetrating brain injury in man afforded the opportunity for such investigations. Various 'primary' visual and tactile capacities were measured, and complex visual and tactile problems were then given.

Problems employing visual material have been used earlier in this laboratory for testing complex processes. These 'complex visual problems' included abstraction- and sorting-tests, matching-from-sample, and conditional reactions.³ The tests were designed to minimize the effect of aphasic symptoms by using procedures approximating those used in animal laboratories, *i.e.* a series of non-verbal differential reactions among alternative stimulus-objects was required. *S* merely had to point at his choice in each of a number of trials, and solution of a problem was inferred from his performance rather than from verbal statement. Relatively little instruction of *S* was needed. Preliminary analysis of the results of some of these tests indicated that performance was particularly vulnerable to lesions in the posterior third of the brain.⁴ Since this part includes the visual receptive area, as well as the so-called 'visual associative' area, the question of whether the defect is restricted to commerce with visual materials, or whether it is more inclusive, has special significance.

We therefore designed tests employing tactual materials which were intended to

³ D. A. Grant and E. A. Berg, A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem, *J. Exper. Psychol.*, 38, 1948, 404-411; Goldstein and Scheerer, *op. cit.*, H. F. Harlow, Responses by rhesus monkeys to stimuli having multiple-sign values, *Studies in Personality*, Q. McNemar and M. A. Merrill (Eds.), 1942, 105-123; Harlow, Solution by rhesus monkeys of a problem involving the Weigl principle using the matching-from-sample method, *J. Comp. Psychol.*, 35, 1943, 217-227; K. S. Lashley, The mechanism of vision: XV. Preliminary studies of the rat's capacity for detail vision, *J. Gen. Psychol.*, 18, 1938, 123-193; Conditional reactions in the rat, *J. Psychol.*, 6, 1938, 311-324; H. W. Nissen, J. S. Blum (Josephine Semmes), and R. A. Blum, Analysis of matching behavior in chimpanzee, *J. Comp. & Physiol. Psychol.*, 41, 1948, 62-74; Conditional matching behavior in chimpanzee; implications for the comparative study of intelligence, *J. Comp. & Physiol. Psychol.*, 42, 1949, 339-356; E. Weigl, On the psychology of so-called processes of abstraction, *J. Abnorm. & Soc. Psychol.*, 36, 1941, 3-33.

⁴ Teuber, W. S. Battersby, and Bender, Performance of complex visual tasks after cerebral lesions, *J. Nerv. & Ment. Dis.*, 114, 1951, 413-429; Effects of cerebral lesions on intellectual functioning in man, *Fed. Proc.*, 11, 1952, 161; Teuber, Some observations on the organization of higher functions after penetrating brain injury in man, *The Biology of Mental Health and Disease*, 1952, 259-262.

be logically analogous to the visual tests already in use. These tactual tests were given to Ss from the same populations (experimental and control) as those from which the Ss for the visual tests were drawn. In this way, we attempted to determine whether performance on complex tactual tasks is as vulnerable to brain injury as is performance on complex visual tasks, and whether complex visual and tactual performance are independently susceptible to injury in different regions of the brain. The majority of the Ss took both the visual and the tactual tests. Comparison of performance on the two series thus required a balanced order of administration, some Ss taking the visual tests prior to the tactual ones, and others taking the two series in the reverse order. This design allowed us also to obtain data on transfer between the two series, and thereby to discover the effect of brain injury (more specifically, of brain injury in a given locus), upon tactual-visual and visual-tactual transfer.

This report presents the results of the tactual series and attempts to evaluate the effect of locus of brain injury upon performance. A forthcoming publication will relate the results reported here to those of the analogous visual tests. Subsequent papers will deal with the brain-injured group subdivided by symptoms rather than by locus of injury, and with the relation of performance on these tests to other measures of tactual perception, and to general intelligence as traditionally assessed.

MATERIAL AND METHODS

Subjects. The Ss were drawn from the case-file of the Psychophysiological Laboratory of New York University College of Medicine. The group consists of men, veterans primarily of World War II, who have sustained penetrating injury to the brain or to peripheral nerves. The establishment of this case-file, and the relative advantages of studying these cases of brain wounds rather than other types of cerebral involvement, have been discussed in previous publications from this laboratory.⁵ For the purpose of longitudinal studies, casualties of the Korean campaign have been added; data from a few of these men increase the heterogeneity of the present sample. The tests were thus given from 1 to 10 yr. after injury, the modal period being about 7 yr.; the resultant symptoms, which ranged from minimal to severe, had become relatively stable in the majority of cases.

The experimental group ($N = 74$) was drawn from the brain-injured population. The control group ($N = 45$) consisted of veterans with injuries to peripheral nerves of the arm or leg. The general comparison was between the performances of these two groups. Further comparisons were made of the various sub-groups of the brain-injured Ss, with each other and with the controls, as indicated in the following paragraphs.

In the brain-injured group, the fact of penetration into the cerebral substance was established by review of roentgenographic and surgical records. The evidence was not sufficient to warrant a distinction between superficial and deep injury, nor could we venture to estimate relative extent of lesion. We did, however, infer the ap-

⁵ Teuber, Battersby, and Bender. *op. cit.* 413-429.

proximate locus of lesion in the forebrain from the wound of entry, from retained foreign bodies, and from reports of observations made at debridement. Table I shows the distribution of the experimental group by locus of lesion.

For the purpose of comparison with the results of the visual tests already reported, the total brain-injured sample was subdivided into groups on the basis of antero-posterior locus of lesion, *i.e.* whether the anterior, intermediate, or posterior sector was damaged. The present classification differs from the preceding one, however, in that the *Ss* were *not* separated into three mutually exclusive categories, according to which sector was predominantly involved. For the study here reported, an *S* was classified as belonging to a given lesion-group if there was *any* damage to the sector in question. Since many of the lesions were not located within one sector, this scheme resulted in groups which were not mutually exclusive. Because of this

TABLE I
DISTRIBUTION OF 74 BRAIN-INJURED *Ss* BY LOCUS OF LESION*

Groups	Bilateral	Unilateral	Left	Right
Anterior	15	17	23	24
Non-anterior	10	32	26	26
Intermediate	13	28	25	29
Non-intermediate	12	21	25	20
Posterior	13	29	30	25
Non-posterior	12	20	20	24
Frontal	17	20	28	26
Non-frontal	8	29	22	23
Temporal	10	19	19	20
Non-temporal	15	30	31	29
Parietal	12	25	25	24
Non-parietal	13	24	25	25
Occipital	10	16	15	21
Non-occipital	15	33	32	31

* It should be noted that each grouping was made independently of the others. For example, although 10 of the *Ss* in the Temporal group were classified as Bilateral, they did not necessarily have lesions to both temporal lobes. The contralateral lesion might be to any of the four lobes.

overlap, these groups were not compared directly with each other. Instead, each was compared with a non-overlapping group composed of all brain-injured *Ss* without damage to the given sector.

Other analyses were made in terms of the lobe of the brain which was damaged, involvement of one or both hemispheres, and laterality of lesion. The principle of grouping for each analysis was analogous to that outlined in the preceding paragraph. Since an *S* simultaneously represented many parameters, it was not possible to equalize the number of *Ss* within each group. A random sample of maximal size for each comparison was therefore drawn from each subgroup.

Comparisons were made by means of analyses of variance. The following analyses were performed:

(I) Overall comparison: Controls vs. Brain-Injured. (II) Analyses by antero-posterior locus of lesion: (A) Controls vs. Anteriors vs. Non-Anteriors; (B) Controls vs. Intermediates vs. Non-Intermediates; (C) Controls vs. Posteriors vs. Non-

Posteriors. (III) Analyses by lobe injured: (A) Controls vs. Frontals vs. Non-Frontals; (B) Controls vs. Parietals vs. Non-Parietals; (C) Controls vs. Temporals vs. Non-Temporals; (D) Controls vs. Occipitals vs. Non-Occipitals. (IV) Analyses by laterality of lesion: (A) Controls vs. Lefts vs. Non-Lefts; (B) Controls vs. Rights vs. Non-Rights. (V) Analysis by involvement of one or both hemispheres: Controls vs. Bilaterals vs. Unilaterals.

Grouping according to antero-posterior locus of lesion. The lesions were plotted on diagrams of the lateral views of the cerebral hemispheres. A line drawn at the level of the greatest longitudinal linear extent of the brain was divided into thirds, and perpendiculars were erected at the points of trisection. The brain was thus divided into anterior (*a*), intermediate (*i*), and posterior (*p*) sectors. An *S* was designated as an *a* if his lesion was completely within the anterior portion, as an *i* if his lesion was completely within the intermediate portion, and as a *p* if his lesion was completely within the posterior portion. He was called an *ai* if his lesion overlapped the boundary between the *a* and *i* subdivisions, and an *ip* if his lesion overlapped the boundary between *i* and *p* subdivisions. If the lesion extended into all three sectors, he was called an *aip*.

All *Ss* with injuries to the anterior sector (*a*, *ai*, and *aip*) collectively constituted the Anterior Group. All brain-injured *Ss* without anterior lesion made up the Non-Anterior Group (*i*, *ip*, and *p*). All *Ss* with injuries to the intermediate sector (*i*, *ai*, *ip*, and *aip*) collectively constituted the Intermediate Group. The Non-Intermediate Group was made up of all brain-injured *Ss* without intermediate lesion (*a* and *p*). All *Ss* with injuries to the posterior sector (*p*, *ip*, and *aip*) collectively constituted the Posterior Group. The Non-Posterior Group was composed of all *Ss* without injury to the posterior sector (*a*, *ai*, and *i*).

Grouping according to lobe injured. Subdivision of the brain into lobes was made according to conventional diagrams. The principle followed in this and subsequent groupings was analogous to that outlined in the preceding section. An *S* was designated an *F* if his lesion fell completely within the frontal lobes, a *T* if his lesion was completely within the temporal lobes, a *P* if his lesion was completely within the parietal lobes, and an *O* if his lesion was completely within the occipital lobes. He was called an *FT*, if his lesion overlapped the boundary between frontal and temporal lobes, and the letter-designations were combined appropriately to represent other lesions encroaching upon more than one lobe. The Frontal Group was made up of *F*, *FT*, *FP*, *FTP*, and *FTPO*—although other combinations are theoretically possible, they did not occur in our sample. The Non-Frontal Group was composed of *T*, *TP*, *P*, *TO*, *PO*, and *PTO*. The Parietal and Non-Parietal Groups, the Temporal and Non-Temporal Groups, and the Occipital and Non-Occipital Groups were constituted in an analogous way.

Grouping according to involvement of one or both hemispheres. *Ss* with lesions to the right hemisphere alone were designated *r*, those with lesions to the left hemisphere alone were called *l*, and those with bilateral lesions were called *rl*. The Unilateral Group was made up of *r* and *l*, and the Bilateral Group of *rl*.

Grouping according to laterality of lesion. The Right Group was composed of all *Ss* whose right hemisphere was damaged (*r* and *rl*), and the Non-Right Group was composed of all brain-injured without right hemisphere involvement (*l*). The Left Group was made up of *l* and *rl*; the Non-Left Group of *r*.

Conditions. The tactual tests were given to an S under one of two different conditions selected in a random fashion. In *Condition I*, the Ss were naïve with respect to the kind of problems they were required to solve, *i.e.* they had not been given the analogous complex visual tests. In *Condition II*, the Ss had previously taken the visual tests. The time between administration of the visual and tactual tests varied from within the same day to over a year, but the factor of intervening time was not differential among the various groups. Most Ss were given all five visual tests before the tactual tests, or conversely, all five tactual tests before the visual tests. There were exceptions to this general rule, but the factor of amount of prior experience with the visual problems was also not differential among the groups. In these

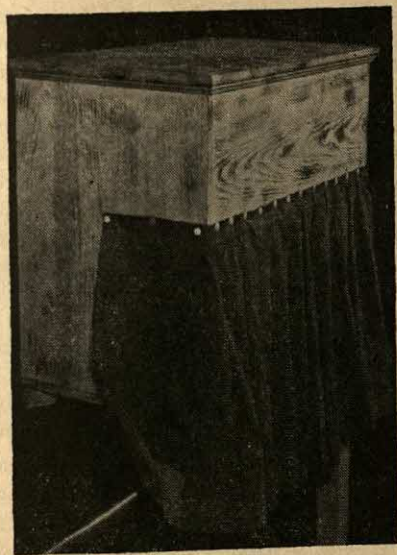


FIG. 1. S'S VIEW OF THE BOX USED IN ALL THE TESTS
S reached beneath and behind the curtain to feel the stimulus-objects
on the floor of the box.

exceptional cases, S was counted as having been tested under Condition II for a particular tactual test if it had been preceded by its visual analogue, and under Condition I for other tactual tests which were not so preceded.

Tests. For all the tactual tests, sight of the test-materials was excluded by placing the stimulus-objects in a wooden box, the front of which was covered by a black cloth curtain (Fig. 1). S reached behind the curtain to feel and manipulate the stimulus-objects.

Each test was composed of a series of trials, a trial consisting of a choice among two or three stimulus-objects, one of which was correct. The position of the correct object varied from right to left in a predetermined order, which was random except that the correct object was not on the same side for more than three consecutive trials. Trials were continued either until the criterion of success (10 consecutive

errorless trials) was met, or until a specified number had been given. In general, *S* was informed only that he must make a choice between the objects in the box, that he would be told after each choice whether he had chosen correctly, and that when he 'got the idea' he could select the correct object every time. If *S* failed to achieve the criterion, the principle by which the correct object on a given test could always be selected was explained to him. Care was taken that he understood the explanation by having him apply the principle on several additional trials.

A description of the tactual tests follows. These five tests were, in the majority of cases, given in the order in which they are described.

Abstraction. This test was composed of three subtests, each consisting of 10 trials. In each subtest, *S* had to discover a common quality among the 10 correct objects which differentiated them from the incorrect ones.

On the first subtest, each trial offered a choice between two common objects, one of which was longer than the other. *S*'s responses were correct if he selected



FIG. 2. EXAMPLE OF ONE TRIAL IN THE ABSTRACTION-TEST
The correct response is to choose the longer of the two objects,
in this case the pencil.

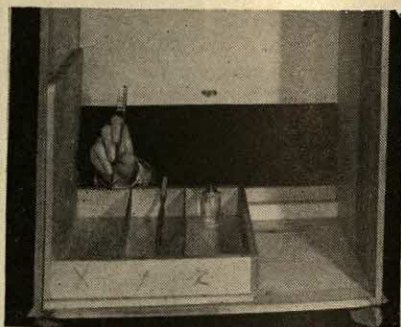


FIG. 3. ONE TRIAL IN THE GENERALIZED MATCHING-TEST
The correct object is identical with the middle object—the sample.

the longer one. The longer objects (*e.g.* pencil, toothbrush, knitting needle, knife) varied from approximately 10–20 cm. in length, and the shorter ones (*e.g.* spool, cork, key, thimble) varied from 1.3 to 3.8 cm. in length. Fig. 2 shows an *S* being given this test. The second subtest required *S* to select soft objects (*e.g.* powder puff, rubber ball, sponge, piece of fur) in contrast to hard objects (*e.g.* coin, pencil, nail, wooden bead). The third subtest opposed metal objects (*e.g.* nail, key, spoon, coil of wire) and objects made of other materials (*e.g.* cork, bottle, piece of cloth, plastic fork); the metal objects were correct. The instructions to the *Ss* were as follows:

Put both your hands behind the curtain into the box. I will give you two objects at a time, one in each hand. One of the objects from each pair is correct, and the other incorrect. All the correct objects are alike in some way. I want you to find out how to select the correct object every time. Just guess on the first trial. After you find out which object of the first pair is correct, pick things that are similar to it in the subsequent pairs. You will be told after each choice whether you were right or wrong.

At the end of each of the subtests, *S* was asked to describe the basis on which his selections had been made. If he had not discovered the 'correct principle, it was explained to him. The instructions and procedure for the second and third subtests were similar. The visual analogue of the tactual abstraction test was the visual abstraction test, previously referred to as 'equivalent and nonequivalent stimuli'.⁶

Generalized matching-from-sample. Each trial consisted of a choice of one among three objects, two of which were exactly alike, and the other different from the pair. The objects used were pencils, bottles, corks, paper clips, locks, keys, nails, tooth-brushes, sponges, and beads. The alternatives in each trial might be two pencils and a bottle, two bottles and a pencil, or other similar combinations. Each object was placed in a separate compartment of a divided tray. One of the pair of identical objects was always placed in the middle compartment; it served as the sample, and the object which matched it was the correct one. If *S* chose the middle object, or the odd object, the response was incorrect. Fig. 3 shows an *S* taking this test. The test was terminated when *S* reached the criterion of 10 consecutively correct responses, or, failing that, after the first error following 39 trials.

For this test *S* used only one hand. In the brain-injured group, the hand opposite the cerebral lesion (or the major lesion, if both hemispheres were involved) was used.⁷ In the control group, either hand was used. The instructions were essentially as follows:

Put your right (or left) hand into the box. On the floor of the box there are three compartments with an object in each. One of the objects is the correct one, and the other two are incorrect. I want you to find out how to select the correct one out of each set of three objects you will be given. You will be told after each choice whether you were right or wrong.

The visual analogue of the tactual Generalized Matching test was the visual achromatic matching test (described in a forthcoming paper).

Conditional matching-from-sample. As in the preceding test, each trial consisted of a choice of one among three objects. The sample (middle) objects, however, resembled *both* of the other objects; it was like one of them in surface texture, and like the other in weight. *S*'s responses were correct if he shifted from one to the other basis of matching in accord with the conditional cue (size of the objects), *i.e.* when the discriminanda were large the correct choice was the object resembling the sample in surface texture, and when the discriminanda were small the correct choice was the object resembling the sample in weight.

The test-objects were two sets of bottles, large and small. Each set consisted of a rough, heavy bottle; a smooth, heavy bottle; a rough, light bottle; and a smooth, light bottle. The large bottles were 12.1 cm. high and 5.1 cm. in diameter; the small bottles were 7.6 cm. high and 3.8 cm. in diameter. A rough surface was obtained⁸ by covering the sides of the bottles with sandpaper (80 grains to the linear

⁶ Teuber, Battersby, and Bender, *op. cit.*, 413-429.

⁷ There were nine cases of hemiplegia included in the brain-injured sample. These *Ss* used the unaffected, or less affected, hand. Their scores will be omitted when the data are analyzed by symptom-groups; they are included here because the presence or absence of sensory or motor defect was not a major consideration in the present analysis by locus of lesion.

inch); the smooth surface was the uncovered glass of the bottle. The heavy bottles weighed 535 grm., and the light bottles weighed 170 grm.

The bottles used on a single trial were all taken from the same set (*i.e.* either all three were large or all three were small). In Trials 1-8, only the large bottles were used; in Trials 9-16, only the small bottles; in Trials 17-24, only the large ones again; in Trials 25-32, only the small ones again. Trials with the large bottles or with the small bottles were given in random order for the remaining eight trials. Sample trials with the large bottles and with the small bottles are shown in Fig. 4.

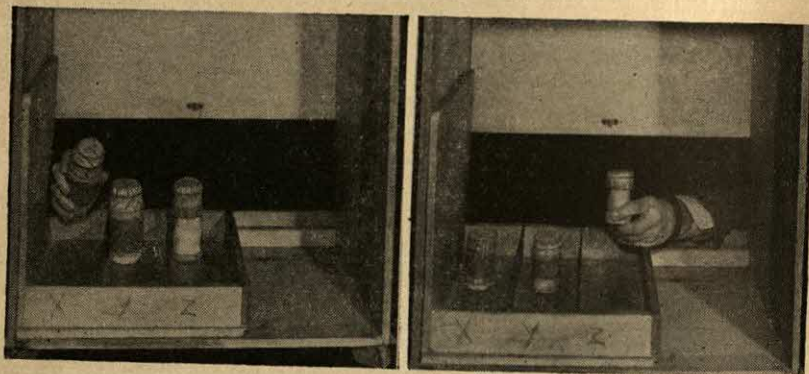


FIG. 4. CONDITIONAL MATCHING-TEST

When the large bottles are presented, the correct response is to choose the one resembling the middle object in surface texture. When the small bottles are presented, *S* must choose the one resembling the middle object in weight.

All 40 trials were given to all *Ss*, regardless of level of performance. The *S* used only one hand to palpate the stimulus-objects. The hand to be used was determined as previously described for the generalized matching test. The instructions were essentially as follows:

Put your right (or left) hand into the box. You will feel the three compartments again, with an object in each. As in the last test, one of the objects is correct, and the other two are incorrect. The middle object is never the correct one, but it has something to do with the solution to the problem. So, on each trial you must pick up and feel all three objects carefully before you make your choice. You will be told after each choice whether you were right or wrong.

The visual analogue of this test was the chromatic matching test (described in a forthcoming report).

Conditional reaction. This problem required a choice between two stimulus-objects, the correct one varying from trial to trial in accord with the conditional cue. The stimulus-objects were a cork (4.4 cm. high, tapering from 2.5 to 3.2 cm. in diameter) and a wooden cube (3.5 cm. on a side). The conditional cue was the orientation of an aperture through which *S* reached into the box containing the stimulus-objects. If the long axis of the aperture was horizontal, the cork was correct (Fig. 5). For this test the front of the box was fitted with a removable wooden wall, in which there was a square window. This window could be covered by either

one of the two sliding panels, one of which contained the horizontally-oriented aperture, and the other the vertically-oriented aperture. For most *Ss*, an aperture of 14 cm. by 5.1 cm. was adequate to enforce proper orientation of the wrist. In a few instances, where the *S's* wrist was unusually large or small, the width of the aperture was adjusted appropriately. The stimulus-objects were fastened to a small turntable, in order to facilitate changing their position from side to side.

The test consisted of 60 trials; it was terminated when *S* reached the criterion, or at the first error after 59 trials. The *S* used only one hand (the same hand as used

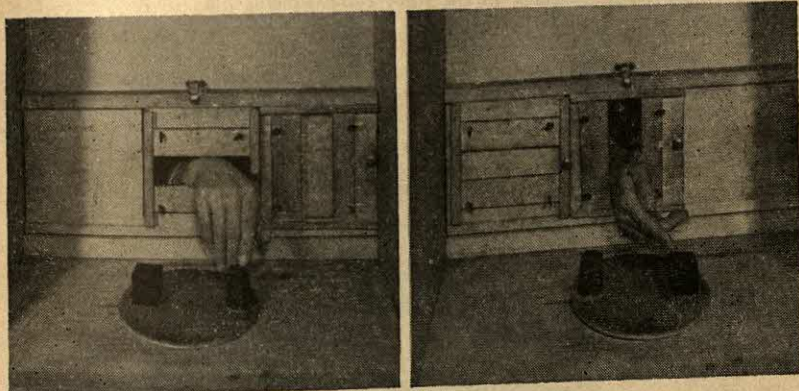


FIG. 5. CONDITIONAL REACTION-TEST

When the aperture through which *S* reaches is horizontally oriented, the correct choice is the cork. When the aperture is vertically oriented, the cube is correct.

in the preceding two tests) to touch and manipulate the stimulus-objects. The instructions were essentially as follows:

Put your right (or left) hand into the box. You will feel two objects on the floor of the box. On each trial, one of them is correct and the other incorrect. I want you to find out how to select the correct object every time. You will be told after each choice whether you were right or wrong.

The visual analogue of this test was the visual conditional reaction (described in a forthcoming report).

Sorting. This test required *S* to sort objects on the basis of kind, number, and length, and to shift from one to another principle of classification as *E* indicated. The objects used were pieces of wire, strips of cardboard, and pencils. One, two, or three objects of a given kind might be presented for *S* to classify. Three lengths of each of the objects were used, 5.1, 9.5, and 19.4 cm.

S was first allowed to examine tactually the three 'models,' each model differing from the other two in the kind of object used therein, in number of objects, and in length of objects. The first model consisted of one short piece of wire; the second, of two medium-sized strips of cardboard; and the third, of three long pencils. The test consisted of 42 trials, on each of which *S* was given an object (or two or three identical objects) and asked to indicate with which of the models it belonged. On Trials 1-7 and 29-35, the correct response was to sort the objects

according to their kind, *i.e.* to place wire with wire, cardboard with cardboard, and pencil with pencil, disregarding number and length. On Trials 8-14 and 22-28, the correct response was to sort objects according to their number, *i.e.* to place groups of objects equal in number together, disregarding their kind and length. On Trials 15-21 and 36-42, the correct response was to sort objects according to their length, *i.e.* to place short objects with short objects, medium-sized objects with medium-sized ones, and long objects with long ones, disregarding their kind and number. In Fig. 6, *S* is sorting according to length.

The models were presented in the three compartments of a divided tray. They remained in place during the entire test. *S* felt the models with one hand and

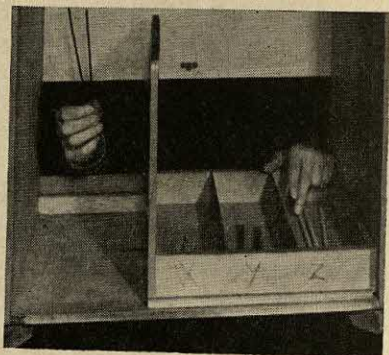


FIG. 6. SORTING-TEST
The *S* is shown sorting according to length.

received the objects to be classified with the other. The hand receiving the test-objects was the same as that used in the preceding three tests. The instructions were essentially as follows:

Put both your hands into the box. With your right (or left) hand, you will feel three compartments with some objects in them. After you have examined these objects, you will be given other objects in your left (or right) hand, and asked to indicate in which compartment they might belong. You might put them in one compartment for one reason, and into another compartment for an equally good reason. For example, if you were sorting a deck of playing cards, you might put the ace of spades with other black cards, with other spades, or with other aces. All these ways of sorting are reasonable. Similarly, on this test, there are various reasonable ways of sorting the objects I am going to give you. At first, I want you to sort them in only one particular way. Whenever you sort them in this way, I will say you are right; if you sort them in any other way, even though it is equally reasonable, I will say you are wrong. Later on, I will change the correct way without warning, that is, I will then say you are wrong when you sort in the way which was previously correct. The correct way of sorting will change several times during the test. Find out as quickly as you can how I want you to sort, and stick to that way until I say you are wrong. As soon as you make an error, shift to some other way of sorting.

The visual analogue of the tactual sorting test was the modified Wisconsin sorting test.

RESULTS

General comparison. The series of tactual tests differentiated between Control and Brain-Injured groups beyond the $1/10$ -% level of confidence. In Figs. 6 through 11 the scores of these groups (and of the various subgroups of the brain-injured sample) for each test and condition are shown. Table II summarizes the analyses of variance. Condition II (in which the Ss had previous experience with analogous tests employing visual materials) yielded a significantly higher score than Condition I (naïve Ss), thus indicating positive transfer. The difference between tests was also highly significant, but there was no evidence in this analysis that the discriminative power of the tests differed one from another or

TABLE II
SUMMARY OF ANALYSES OF VARIANCE

Each column shows the values of *F* derived from a separate analysis. 1, Controls vs. Brain-Injured; 2, Controls vs. Anteriors vs. Non-Anteriors; 3, Controls vs. Intermediates vs. Non-Intermediates; 4, Controls vs. Posteriors vs. Non-Posteriors; 5, Controls vs. Frontals vs. Non-Frontals; 6, Controls vs. Parietals vs. Non-Parietals; 7, Controls vs. Temporals vs. Non-Temporals; 8, Controls vs. Occipitals vs. Non-Occipitals; 9, Controls vs. Bilaterals vs. Unilaterals; 10, Controls vs. Lefts vs. Non-Lefts; 11, Controls vs. Rights vs. Non-Rights.

Source	1	2	3	4	5	6	7	8	9	10	11
Between groups	19.42*	5.18†	6.36†	6.14†	6.92*	7.27*	9.19*	3.28†	3.59†	3.69†	5.36†
Between conditions	7.67†	5.86†	7.56†	7.50†	7.33†	14.18*	10.03†	8.79†	7.04†	6.15†	10.48†
Between tests	20.65*	24.86*	26.23*	38.39*	31.79*	31.50*	19.64*	26.34*	26.20*	26.28*	21.12*
Groups×conditions	1.95	1.27	3.66†	1.82	0.40	3.06†	0.44	0.08	0.65	1.63	2.15
Groups×tests	2.16	0.95	0.50	0.59	0.85	0.37	1.48	1.52	0.88	0.20	0.87
Conditions×tests	1.35	2.76†	2.50†	2.94†	2.03	3.14†	4.46†	4.94*	2.04	2.33	3.21†
Groups×conditions×tests	1.81	1.20	1.57	1.26	2.09†	1.18	0.44	1.11	0.54	0.67	0.95

* Significant at the 0.1-% level of confidence.

† Significant at the 1-% level of confidence.

‡ Significant at the 5-% level of confidence.

that the effects of the conditions were differential among the tests. The order of difficulty of the tests, from easiest to most difficult, was Abstraction, Generalized Matching, Conditional Matching and Sorting (for which the mean scores were identical), Conditional Reaction. The mean score for each test (all groups and conditions combined) was significantly different from that of every other, with the exception already noted. Inspection of the range of scores in the various groups indicated that the variances of the brain-injured and control groups were approximately the same. Any slight difference in the variances might decrease the power of the tests to detect differences, but would not vitiate our positive findings.⁸

Analyses of brain-injured subgroups. Analyses by antero-posterior locus

⁸ W. G. Cochran, Some consequences when the assumptions for the analysis of variance are not satisfied, *Biometrics*, 3, 1947, 22-38.

of lesion, by involvement of one or both hemispheres, and by laterality of lesion yielded no significant differences between the brain-injured subgroups. Analyses by lobe injured, however, showed that the group with injury to the temporal lobe made a significantly poorer mean score (5-% level of confidence) than that of the brain-injured group without temporal lesion. No other division of the total brain-injured sample produced a significant difference among the mean total scores of the resulting subgroups. Consideration of the number of analyses made, and the level of significance obtained in this case, suggests that this result should be regarded as hypothetical only.

The performance of the Control Group was significantly superior to those of all brain-injured subgroups with two exceptions: the scores of the Unilateral Group and of the Non-Left Group (Ss with lesion of the right hemisphere alone) did not differ significantly from that of the Controls. The Unilateral Group, however, was not significantly superior in score to the Bilateral Group; nor was the Non-Left Group significantly superior to the Left Group.

Although the main effects (of groups, conditions, tests) were significant in all analyses, the interactions of these variables were in most cases not significant. Only in one analysis could we infer that the brain-injured subgroups benefited unequally from experience with the visual tests. When the total brain-injured sample was divided into subgroups with and without lesion of the parietal lobe, Condition II placed the Parietal Group at a significant disadvantage to both of the other groups, Controls and Non-Parietals ($1/10$ -% level of confidence in the Parietal-Control comparison, and just short of the 2-% level in the Parietal-Non-Parietal comparison). The Parietal Group was not significantly different in score from either of the others in Condition I. Of these three groups, the Parietals alone showed no significant elevation of scores after experience with the visual tests.

When Conditions I and II were combined, no test was significantly more effective than any other in differentiating the groups. The interaction of groups, conditions, and tests was, however, significant in the analysis of Controls vs. Frontals vs. Non-Frontals. Further examination of the individual tests under each condition showed that the group with lesion of the frontal lobe was significantly inferior to the brain-injured group without frontal lesion (and to the Controls) on Conditional Matching in Condition II. No other significant differences among brain-injured subgroups were found.

DISCUSSION

This study explores the differential effects of locus of brain injury upon solution of problems perceived through the sense of touch. The tactual tests were designed to compare the results, in brain-injured and normal groups, with those obtained on an approximately parallel series of visual problems. The finding of significant improvement in score on the tactual tests after experience with the visual tests suggests that the two series were perceived as similar by the Ss. The transfer from one modality to the other was, however, far from perfect, indicating that the perceptual similarity of the tests was not as great as the similarity of their logical structures would lead one to expect.

The series of tactual tests differentiated the total brain-injured group from the controls at a high level of confidence. In search of a more specific localization of the deficit in the brain-injured group, we divided the group in several ways: by antero-posterior locus of lesion (anterior, intermediate, posterior), by lobe injured (frontal, parietal, temporal, occipital), by involvement of one or both hemispheres (bilateral, unilateral), and by laterality of lesion (left, right).

Preliminary analyses of the results of the analogous visual tests had indicated that "patients with posterior lobe lesions exhibit the poorest performance, those with anterior penetrations are somewhat better, and those with intermediate lesions do best."⁹ The results of the tactual series, in contrast, show no hierarchy of performance in terms of antero-posterior locus of lesion. The controls were significantly superior to all such subgroups of the brain-injured, and no subgroup was significantly inferior to any other.

Analyses by lobe injured, however, suggested that lesion of the temporal lobe may produce poorer scores than other cerebral injuries. These results recall earlier findings in monkeys which imputed some rôle in tactual function to the temporal lobes.¹⁰ In the complex tactual tasks as described here for man, temporal lesion was not, however, the only factor making for poor performance. The group without temporal lesion was also significantly inferior to the controls.

Analyses by involvement of one or both hemispheres and by laterality of lesion suggest other relevant variables. Neither the total group with

⁹ Teuber, *op. cit.*, 259-262.

¹⁰ J. S. Blum, Cortical organization in somesthesia, *Comp. Psychol. Monog.*, 20, 1951, 219-249; J. S. Blum, K. L. Chow, and K. H. Pribram, A behavioral analysis of the organization of the parieto-temporo-preoccipital region, *J. Comp. Neur.*, 93, 1950, 53-100.

unilateral lesion nor the group with lesion to the right hemisphere alone had mean scores significantly different from the controls. Deficit cannot be ascribed wholly to bilaterality or to left-sided lesion, however, because neither of these lesion-groups differed significantly from the remainder of the brain-injured Ss. Inspection of the composition of the Temporal and Non-Temporal groups shows that the effects of laterality and of bilaterality were not confounded with those of lobe. Moreover, although we made no statistical comparison of the means of unilateral right and unilateral left temporal lesion-groups, it is of interest that the means of these groups differed by only one point.

The relation of our findings to those of other investigators who have studied the effects of temporal lobe lesion in man is uncertain. To our knowledge, the only tactual problem heretofore used in such investigations is the formboard. Hebb described a patient with right temporal lobectomy to whom he administered this test for the purpose of comparison with performance (non-language tests) based upon vision.¹¹ He noted that "the patient was if anything worse in non-visual than in visual performance, by comparison with normal subjects." Recently, Milner¹² and Teuber and Weinstein¹³ have duplicated Hebb's results with a large number of cases, but have further shown that this defect does not occur after lesions of the left temporal region.

Riopelle, Alper, Strong, and Ades,¹⁴ using a series of visual discrimination tasks, found that monkeys with lesions of the temporal lobes are deficient in the formation of 'learning sets'.¹⁵ The question might therefore arise as to whether the low mean score of our groups with temporal lesion was dependent particularly on lack of transfer from the visual tests. It can be seen from Fig. 7, however, that this group appears at a somewhat greater disadvantage on the tactual tests when they *precede*, rather than follow, the visual tests. Students of learning sets have confined their attention to interproblem effects within the visual modality; investigations on formation of intermodal learning sets might be a profitable extension of these experiments. The present study afforded the

¹¹ D. O. Hebb, Intelligence in man after large removals of cerebral tissue; defects following right temporal lobectomy, *J. Gen. Psychol.*, 21, 1939, 437-446.

¹² Brenda Milner, Intellectual effects of temporal lobe damage in man, Unpublished Ph.D. dissertation, McGill University, 1952.

¹³ Personal communication.

¹⁴ A. J. Riopelle, R. G. Alper, P. N. Strong, and H. W. Ades, Multiple discrimination and patterned string performance of normal and temporal-lobectomized monkeys, *J. Comp. & Physiol. Psychol.*, 46, 1953, 145-149.

¹⁵ H. F. Harlow, The formation of learning sets, *Psychol. Rev.*, 56, 1949, 51-65.

opportunity to investigate intermodal transfer in normal and brain-injured groups.

Although transfer from one modality to the other was not measured directly, differential effects of the order of administration of the tests could be inferred when the interaction of groups and conditions was significant in the analyses of variance. Since this interaction was not significant in our general analysis, we cannot conclude that the brain-injured

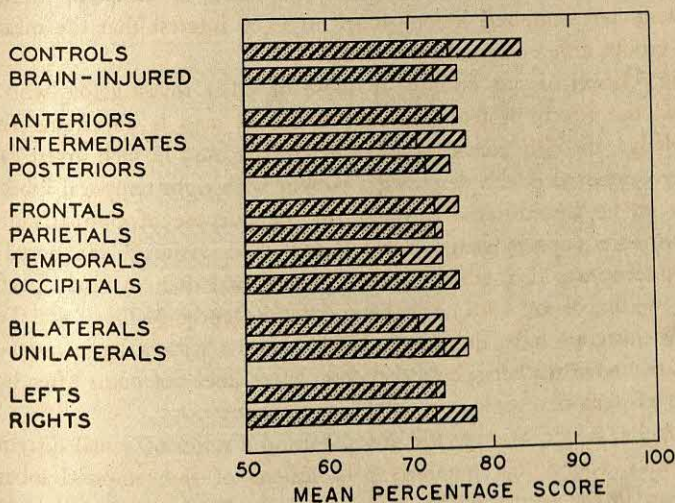


FIG. 7. MEAN PERCENTAGE SCORES ON ALL TESTS COMBINED FOR THE CONTROLS, THE TOTAL BRAIN-INJURED GROUP, AND THE VARIOUS BRAIN-INJURED SUBGROUPS INDICATED

Before averaging, the scores of all tests were converted to their proportional values on a common scale, with 50% representing a chance score and 100% representing errorless performance. The score for Condition I is shown by stippling, and that for Condition II by cross hatching.

group as a whole benefited less than the controls from experience with the visual tests. Analyses of brain-injured subgroups, revealed, however, that the group with lesion to the parietal lobe, in contrast to the controls and to the remainder of the brain-injured group, showed no significant elevation of scores after having taken the visual tests.¹⁶ Consistent with

¹⁶ It may be seen in Fig. 6 that the absolute amount of improvement was similar in the Parietal group and in the Left group. The lack of significance of the group-condition interaction in the analysis of Controls versus Lefts versus Non-Lefts does not, however, allow the inference of differential transfer among these groups. It is possible that the effects of parietal and of left-sided injury interact, *i.e.* that only Ss with injury to the left parietal lobe have especial difficulty in transfer. Inspec-

this defect in transfer was the finding that the group with parietal lesion was significantly inferior not only to the controls but also to the complementary brain-injured group in Condition II (visual tests given prior to tactual tests). Comparison of the amounts of improvement in the subgroups with and without lesion of the intermediate sector (a division which splits the parietal lobe) suggests that the posterior parietal region is of greater importance than the anterior, for transfer in the Intermediate group was considerable and approached significance. Whether the defect

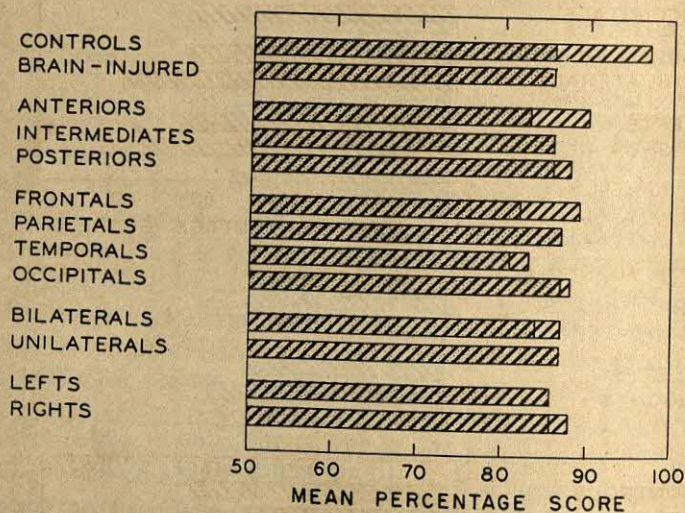


FIG. 8. MEAN PERCENTAGE SCORES (CONVERTED) ON THE ABSTRACTION-TEST

shown by the Parietal Group reflects inadequacy of concepts derived from visual experience, or difficulty with their application to tactual materials, remains a problem. Reanalysis of the visual tests, with similar subgroupings to those used here, may clarify the issue. Discussion of evidence for relative vulnerability of visual and tactual performance to brain injury in man likewise will be postponed until additional information on the visual tests becomes available.

A final comment should be made with regard to subdivision of the brain-injured sample by locus of lesion. Grouping according to antero-posterior locus, as used in this laboratory, has proved useful in revealing significant differences between Ss with field-defects (posterior lesions) and

tion of the means for each test under Condition II in unilateral left and unilateral right parietal lesion-groups suggests that this might be the case. Examination of a larger number of Ss would be necessary, however, in order to establish this point.

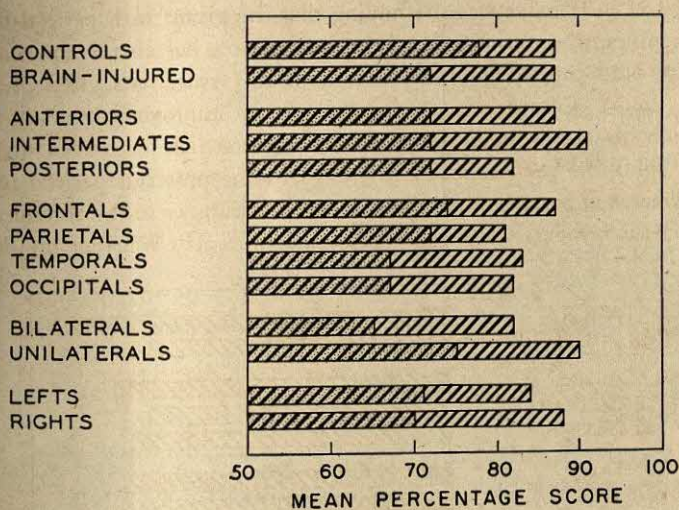


FIG. 9. MEAN PERCENTAGE SCORES (CONVERTED) ON THE GENERALIZED MATCHING-TEST

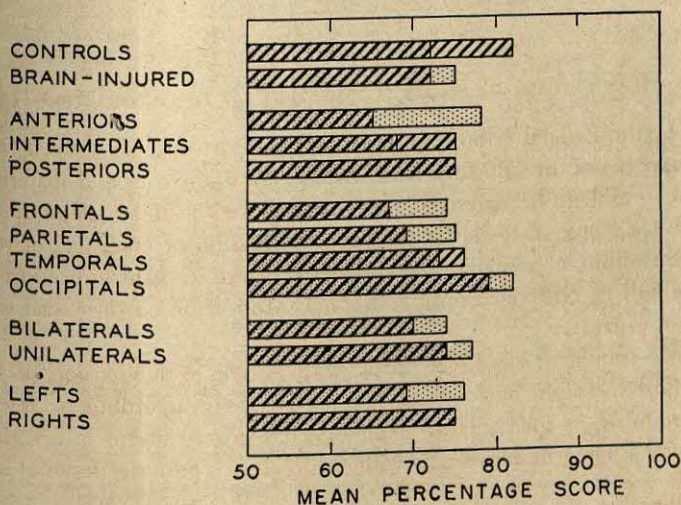


FIG. 10. MEAN PERCENTAGE SCORES (CONVERTED) ON THE CONDITIONAL MATCHING-TEST

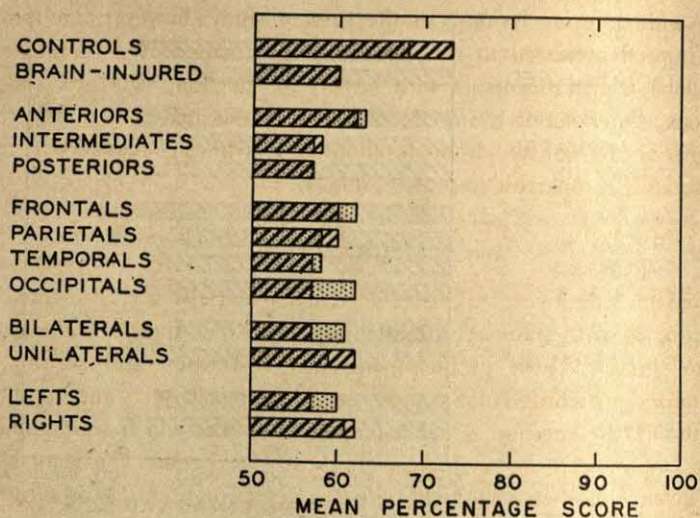


FIG. 11. MEAN PERCENTAGE SCORES (CONVERTED) ON THE CONDITIONAL REACTION-TEST

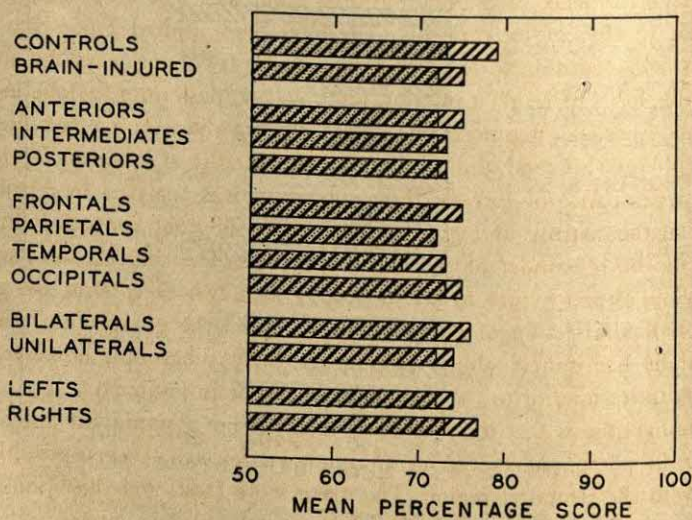


FIG. 12. MEAN PERCENTAGE SCORES (CONVERTED) ON THE SORTING-TEST

those with anterior lesions. In the present study, however, grouping by lobe injured appeared to generate more information. We recognize that each lobe is heterogeneous with respect to function, structure, and connections. Finer subdivisions are unsuited to our material, however, and designation by lobes, although admittedly artificial, makes our results more readily comparable to those of others.

SUMMARY

Problems employing exclusively tactual materials were presented to a group of Ss with traumatic brain injury and to an appropriately matched control group. These problems included abstraction and sorting tests, generalized matching-from-sample, conditional matching, and conditional reaction. Their solution depended upon principles which were similar to those employed in an analogous group of visual tests. The tactual series was given either preceding the visual series (Condition I) or subsequent to it (Condition II).

Condition II yielded a significantly higher score than Condition I, thus showing some degree of positive transfer from one modality to the other. The tactual tests differed significantly among themselves in level of difficulty. The order of difficulty, from least to most difficult, was Abstraction, Generalized Matching, Conditional Matching and Sorting (for which the mean scores were identical), Conditional Reaction.

The performance of the brain-injured group as a whole was significantly inferior to that of the controls. The total brain-injured group was then subdivided according to locus of lesion in the following ways: by antero-posterior locus of lesion (anterior, intermediate, posterior) by lobe injured (frontal, parietal, temporal, occipital), by involvement of one or both hemispheres (bilateral, unilateral), and by laterality of lesion (left, right). The performance of each of these subgroups was analyzed in relation to those of the controls and of a complementary subgroup consisting, in each case, of the remainder of the brain-injured Ss. The controls were significantly superior in score to all subgroups with two exceptions, the group with unilateral lesion of either hemisphere and the group with lesion to the right hemisphere alone. Neither of these subgroups, however, was significantly superior to the remainder of the brain-injured Ss.

The hypothesis that locus of brain injury is not a significant differential factor in performance-decrement cannot be rejected, except possibly for injury to the temporal region. The mean score (tests and conditions combined) of the group with injury to the temporal lobe was significantly

lower than that of the brain-injured group without temporal lesion. No other division of the total brain-injured sample produced a significant difference among the resulting subgroups.

Elevation of scores after experience with the analogous visual tests was not significantly less in the brain-injured group as a whole than in the controls. When the total brain-injured sample was divided into groups with and without lesion to the parietal lobe, however, the group with parietal lesion showed no significant transfer effect, in contrast to the controls and to the remainder of the brain-injured. In Condition II the performance of the parietal group was significantly inferior to the controls and to the brain-injured group without parietal lesion. In no other analysis could we infer that the order of administration of the tests had differential effects dependent upon locus of brain lesion.

ON THE COMBINATION OF ASSOCIATIVE PROBABILITIES IN LINGUISTIC CONTEXTS

By DAVIS HOWES, Aero Medical Laboratory, Wright-Patterson AFB,
and CHARLES E. OSGOOD, University of Illinois

It is a commonplace that meanings of words depend upon the contexts in which they occur. This dependence sets a fundamental problem in the psychology of language: calculation of the psychological effects of a word in its context from the individual properties of the word and of the contextual elements. Recognizing a conventional distinction between the linguistic and non-linguistic contexts of a person's speech, we may conveniently subdivide the former into (a) the *homogeneous linguistic context*—the context provided by his own previous language behavior; and (b) the *heterogeneous linguistic context*—that provided by the utterances of other persons in his environment. Recent work by Shannon has provided a statistical model for describing certain problems of homogeneous sequences and has stimulated several experimental studies in that area.¹ In the present paper we shall be concerned solely with heterogeneous linguistic contexts; i.e. with the prediction of the language behavior of an experimental subject from the language behavior of *another* person in his environment.

For experimental purposes we take a sequence of four words spoken by an experimenter (*E*) and measure as a dependent variable the probability that a given word will be emitted as an association to the last word of the sequence. This is a modified form of the conventional word-association experiment. Since the sequence is spoken by *E*, the properties of the words constituting it can be controlled as independent variables. The strengths of the associative effects of each of the first three words of the sequence upon the subject's (*S*'s) response to the fourth word is the property investigated in the following experiments. Three studies will be reported. If we designate the first three words of the four-word sequence as the *context* and the fourth word as the *test-word*, the independent variables defining the three

* Accepted for publication October 23, 1952.

¹ C. E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.*, 27, 1948, 379-423, 623-656; G. A. Miller and J. A. Selfridge, Verbal context and the recall of meaningful material, this JOURNAL, 63, 1950, 176-185; E. B. Newman, The pattern of vowels and consonants in various languages, this JOURNAL, 64, 1951, 369-379; C. E. Shannon, Prediction and entropy of printed English, *Bell Syst. Tech. J.*, 30, 1951, 50-64.

experiments are as follows: (1) *interposition*—the number of stimulus-words separating a given contextual word from the test-word; (2) *density*—the number of contextual words having similar associative effects; (3) *frequency of occurrence*—the number of times a contextual word occurs in a general sample of the language under study. Since the experiments attempt to predict the probability that a particular association will be emitted following stimulation with a sequence of words from knowledge of its probability following single words, these studies take the general form of experiments on the combination of word probabilities.

Definitions and symbols. From the above description it can be seen that the basic idea requiring definition is that of the probability of a word occurring as an association. For the general case we can define the *associative word-probability* $P_{j_1, j_2, \dots, j_m} (i_1, i_2, \dots, i_n)$ for two populations of subjects U and W as the limiting relative frequency with which the sequence of words $i_1, i_2, \dots, i_n$ is emitted by population U following the emission of the sequence of words $j_1, j_2, \dots, j_m$ by population W . To give the definition experimental meaning, certain boundary conditions have to be specified. First, prediction will be attempted for the probability of only the first word of the sequence of associated words, i_1 (which can therefore be written simply as i). Similarly, only sequences of *four* stimulating words j will be considered. The stimulating population W will be a single person (one of the E s), and the associating population U will consist of a class of college students. One other very important condition must be specified viz. the instructions given the S s. These are, briefly, to associate only to the final word of the stimulating sequence. This final word will therefore be called the *test-word* and will be designated j_t . The entire sequence of four stimulating words, called a *word-set*, will be written j_1, j_2, j_3, j_t to indicate the special emphasis placed on the test-word by the instructions. The three words that precede the test-word constitute a *context*. The form of our experiments can thus be represented by the expression $P_{j_1, j_2, j_3, j_t} (i)$. In words, this expression represents the probability that word i will occur as the association to a sequence of four words when the S s are instructed to associate to the last word only.

Associative probabilities following *single-word* stimulation, such as those reported by Kent and Rosanoff,² will be called *first-order* associative probabilities and written $p_j(i)$. Since the strength of these associations is the basic quantity in the present experiments, subscripts will be used to indicate different values of this variable. *Large* first-order associative probabilities will be indicated by *letter* subscripts to the stimulus-symbol, *small* first-order associative probabilities by *numerical* subscripts. Thus $p_{1a}(i) \gg p_{11}(i)$. It will be convenient to define the *associative effect* of a stimulus-word j as the probability of the association i following j in the conventional word-association experiment. In the above inequality, for example, a word represented by j_a (e.g. *man*) can be said to have a stronger associative effect than j_1 (*boy*) upon the association i (*woman*) for the population studied by Kent and Rosanoff, since in their tables the relative frequency with which *woman* is associated to *man* is 0.394 while its relative frequency of association to *boy* is only 0.002. These symbols

² G. H. Kent and A. J. Rosanoff, A study of association in insanity, *Amer. J. Insanity*, 67, 1910, 37-96, 317-390.

for the strength of first-order associative effects will be carried over to the word sets: thus, P_{j_a, j_1, j_2} , $j_t(i)$ indicates that $p_{j_a}(i)$ and $p_{j_t}(i)$ are large relative to $p_{j_1}(i)$ and $p_{j_2}(i)$.

The effect of the context of a word-set upon the probability of the association i will be measured as the difference between the probability of i following the word-set containing that context and the probability of i following the test-word with a control context consisting of nonsense words and three-place numbers (represented by j'). It is assumed that such contextual materials affect the associative response negligibly. Values of $p_{j_t}(i)$, estimated from the Kent-Rosanoff tables, can also be put to this purpose. But the Kent-Rosanoff data can be applied to the present population only with reservation in view of the type of Ss they used and the historical period in which their measurements were made.

The extent to which the word-set constitutes a meaningful sequence that might be emitted in the course of ordinary continuous speech can be expected to influence the associative response. Response to the word-set "*Mary had a little,*" for example can be expected to differ greatly from that to a set, such as "*a had Mary little,*" which presents the same components in a less familiar sequence. To control for this factor we can specify that the *transitional probabilities* for populations U and W shall be negligible where the transitional probability $P^{*j_{k-1}(j_k)}$ is defined as the probability that word j_k will follow a given word j_{k-1} in samples from the continuous discourse of a stated population. Populations U and W in the present case both represent samples of the general population of English-speaking persons of college education.

It should be clear that the symbols defined here represent only experimental variables that we have attempted to control in the present studies, and do not imply that the dependent variable can be completely specified by these concepts. All kinds of complex interactions, in addition to those specified by the transitional probabilities of the word-set, may modify the effects upon the subject of each contextual word. Indeed, one of the purposes of the experiments is to ascertain the extent to which a few simple conditions, such as those defined above, can determine the probability of a given associative response. The elucidation of complex interactions between the effects of different elements of a word-set is a natural corollary of such an experimental program.

All measurements necessary to specify the independent variables of these experiments could be obtained from extensive tables of first-order associative and transitional probabilities analogous to those published for the former by Kent and Rosanoff and others.³ Sufficiently extensive tables, unfortunately, exist in neither case, so that it has been necessary to make subjective estimations of these probabilities. Little validity can be imputed to subjective estimations of fine differences in associative or transitional probabilities; the experimental designs have therefore been simplified to require only these very simple judgments: (1) that a given word is a *highly probable* association, in the conventional word-association test, to a second given word; (2) that a given word is a *highly improbable* association to a second given word; and (3) that a given sequence of words will occur *only very rarely* (if at

³ Kent and Rosanoff, *op. cit.*; J. O'Connor, *Born That Way*, 1928; Herbert Woodrow and F. Lowell, Children's association frequency tables, *Psychol. Monog.*, 22, 1916, No. 97, 1-110.

all) in ordinary continuous discourse, and hence represents very low transitional probabilities. The design of the experiments makes it necessary that the first two judgments be valid only relative to each other; thus only two subjective judgments are really required for the experiments. Selection of sequences with low transitional probabilities can be facilitated by eliminating from the word-sets all connective words (articles, conjunctions, prepositions, etc.), since sequences of high transitional probability usually include one or more of these words.

While no validity studies specific to these judgments have been conducted, the ability of subjects to estimate the relative probabilities of occurrence of words would indicate that such simple judgments can be carried out with small error.⁴ Such estimates, of course, do not possess the objectivity of measurements that are independent of the experimenter's judgment, and the results to be presented below must be interpreted with that limitation in mind. At the same time, it should be emphasized that the difficulties are not inherent in the conceptual design of the investigation, but are artifacts of a temporary lack of empirical information. Whenever more extensive tables of associative and transitional frequencies are available, the present experiments can be replaced by more rigorous ones.

Subjects and instructions. Approximately 200 men, students in the introductory psychology course in Yale College, served as Ss in these experiments. They were divided into four groups,⁵ each of which was given the same test-words but different contexts. The following instructions were read to all the groups.

This is an experiment on what is known as 'free association.' When a person calls out a word and you say (or write) the first word that comes into your head, that is a free association. Thus, for example, if I should say *tree* you might immediately think *trunk* or *leaf* or *green* or *sleep*, or any word whatsoever. The present experiment differs from this simple type in one respect: I will read *four* words slowly and you are to listen carefully to all of them, but you are to free-associate to the last word only. Thus if I should say '*toy, come, wretched, book,*' you would then write down the first word that *book* made you think of. This is important: listen to each word carefully, but respond only to the last one. Be sure to get the first association down. There will be 50 such word-sets altogether. I will call out the number of the set before I read off the words. Although the test-words will all be meaningful, the other three will sometimes be nonsense words or numbers. Try to avoid responding with one of the words in the set, but write it down if it comes up strongly.

Before the experiment was begun the main rules were repeated: "(1) Listen to all four words carefully. (2) Free-associate to the last word only. (3) Put down the first word that comes to mind; write it immediately. (4) Try to avoid writing down one of the words used in the word set." To set the test-word off from the context, *E* paused slightly between reading the last word of the context and the test-word and enunciated the latter with an intonation of finality (the tacit period of a spoken sentence). Each word-set required about 5 sec. to read. The same *E* read the instructions and word-sets to all groups of Ss.

Materials. The first 5 word-sets read to each group were given to show the Ss the procedure. Of the remaining 45 sets, 10 were devoted to Experiment 1 (inter-

⁴ D. Howes, The definition and measurement of word probability, Doctoral Dissertation, Harvard University, 1950, 33-36.

⁵ Scheduling difficulties resulted in a slight imbalance among the groups. The actual number in each group varied from 46 to 57.

position), 10 to Experiment 2 (density), and 8 or 7, depending on the group, to Experiment 3 (frequency of occurrence). The remaining word-sets concerned variables not reported in this paper (the word-sets for these experiments appeared in random order throughout the list of 50).

Since the *Ss* were instructed to avoid using contextual words as associations, it was necessary to exclude from each context any word likely to occur as an associate to the corresponding test-word. To insure this condition, all test-words were taken from the list of 100 stimulus-words for which Kent and Rosanoff published 1000 associations, and no word was admitted to a context if in those tables it appeared as an association to the test-word with a frequency greater than 0.005.

Treatment of data. The dependent variable in these experiments, $P_w(i)$,⁶ is the proportion of *Ss* who respond to a word-set with a particular association, *i*. The associations to be so measured must be selected carefully for the probability with which they occur as first-order associations to the test-word. For if $p_{1t}(i)$ were very high, no context could effect a considerable increase in the probability of *i*, which must always be less than unity; while too low a value of $p_{1t}(i)$ would be impossible to measure accurately with groups of only 50 subjects. If a cluster of associations, all having similar first-order associative relations to the contextual words, is taken as the dependent event *i*, the first-order associative probability of each response-word can be small while the probability of the whole cluster can be large enough to present no serious difficulty of measurement in samples of 50. This solution has been followed in the present experiments. For the word-set: *devil, fearful, sinister, dark*, for example (a set, used in the density experiment, in which each contextual word should have a strong associative effect upon the response), the frequency of occurrence of *i* was tabulated for the associative cluster (*bad, evil, fear, fright, frightening, gloomy, hell, mysterious, scared, scary*).

The chief danger in specifying these clusters is that personal idiosyncrasies in the meanings of words will influence the selection. Several precautions were observed in order to minimize this effect: (1) each *E* examined all 200 associations made to a given test-word without distinguishing among those made under different experimental conditions; (2) each *E* excluded any word about which he felt doubtful (*i.e.* with regard to its associative relations to the contextual words); (3) separate lists of associative clusters were drawn up by the two *Es*, and no important differences appear when results are calculated separately for each list; (4) a final list was prepared, including only words listed by both *Es*, in which a still more strict criterion for selection was adopted. Calculations given below are for this final list. As additional assurance against the problem posed by associations of high probability, any word occurring to the test-word with a relative frequency of 0.02 or higher in the Kent-Rosanoff tables, or of over 0.04 to the control word-sets with nonsense-word contexts, was excluded from the clusters.

While it is unfortunate that these subjective estimations had to be made, it should be remembered that the discrimination that is basic to all these approximations is simply the ability to select one set of words that occur as associations to a given stimulus-word much more frequently than to a second set of words. This discrimination is made especially easy in the present study by the fact that each associative cluster need include only a few words. Thus any word about which there was doubt

⁶ The subscript *w* appears here as a general symbol for word-set.

could be eliminated. As will be seen, the experimental design itself provides a further safeguard: any word selected for an associative cluster is counted in all experimental groups, including the control, and poor selection would tend to minimize effects of the experimental variables rather than to exaggerate them.

Results: Experiment 1: Interposition. It can be expected that an increase in the number of stimulus-words, $j_1, j_2, \dots$ that are presented between a given contextual word j_a and the moment of the S's response will tend to decrease the effect of j_a upon the associative response. The present experiment explores this relation.

The experiment involves measurement of the probability of a given associative cluster i under three conditions: one in which the experimental word j_a is the first word of the context (hence farthest removed from j_t); a second in which it is the second word of the context; and a third in which it is the last word in the context (hence nearest to j_t). The remaining words in the context are neutral with respect to i . A control context made up of 3-place numbers or nonsense words of two or three syllables must be added. The four word-sets of this experiment can thus be represented:

$$\begin{array}{ll} \text{Condition I, } j_a, j_1, j_2, j_t; \\ \text{Condition II, } j_2, j_a, j_1, j_t; \\ \text{Condition III, } j_1, j_2, j_a, j_t; \\ \text{Condition IV, } j'_1, j'_2, j'_3, j_t. \end{array} \quad [1]$$

For example, let $j_a = \text{skin}$, $j_t = \text{rough}$, and consider the association *hands*. The word-sets read to the Ss might then be: I, *skin, hour, utter, rough*; II, *hour, skin, utter, rough*; III, *utter, hour, skin, rough*; IV, *318, hokiba, rafuny, rough*. It is assumed that neither *hour* nor *utter* has an appreciable effect upon the occurrence of *hands* as an association.

The amount of data this experiment generates can be tripled by selecting second and third associations different from the first one. Representing by i_a the first association, which is a strong first-order associate of j_a , we can write the two other associations i_β and i_γ . We now select a word that has a strong first-order associative effect on i_β but negligible first-order associative effects on i_a and i_γ . This word, which can be represented by j_b , can be used in place of the neutral word j_1 of the contexts, and thus its separation from j_t will be varied just as the separation between j_a and j_t is varied. A third word, j_c , can be chosen in like manner for strong first-order associative effect on i_γ and negligible effects upon i_a and i_β . It can replace j_2 in the word-sets of Equation [1]. We can then rewrite the first three word-sets of the experiment:

$$\begin{array}{ll} \text{Condition I, } j_a, j_b, j_c, j_t; \\ \text{Condition II, } j_c, j_a, j_b, j_t; \\ \text{Condition III, } j_b, j_c, j_a, j_t; \end{array} \quad [1a]$$

where it is assumed that $p_{1a}(i_a)$, $p_{1b}(i_\beta)$, and $p_{1c}(i_\gamma)$ are large relative to the remaining associative probabilities, $p_{1a}(i_\beta)$, $p_{1a}(i_\gamma)$, $p_{1b}(i_a)$, $p_{1b}(i_\gamma)$, $p_{1c}(i_a)$ and $p_{1c}(i_\beta)$. For illustration, let $j_t = \text{rough}$, $i_a = \text{hands}$, $i_\beta = \text{storm}$, and $i_\gamma = \text{rocky}$, and let the three contexts of [1a] be: I, *skin, wind, mountain*; II, *mountain, skin, wind*; III, *wind, mountain, skin*. We then determine the combined probability of

hands following both *skin* and *rough* when zero, one, or two neutral contextual words are interposed between them, just as before. In addition, the same calculations can be made for *storm* following *wind* and *rough* and for *rocky* following *mountain* and *rough*. Thus a single word-set provides three determinations of the function.

Fig. 1 shows the probability of associative clusters, $P_w(i)$, as a function of the number of interposed neutral words. Thirty determinations were made for each number of interposed words, three for each of the 10 word-sets. Thus $P_w(i_a)$ is plotted opposite the abscissal number zero for Condition III, in which j_a is the last contextual word, and opposite the number

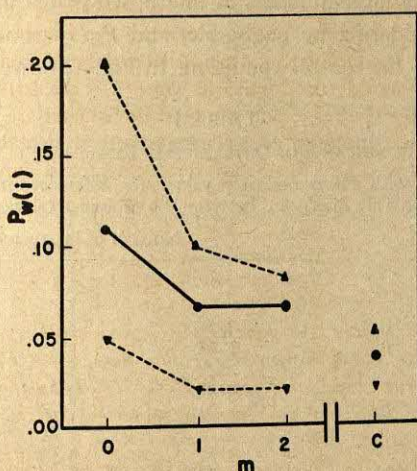


FIG. 1. THE PROBABILITY OF AN ASSOCIATIVE CLUSTER, $P_w(i)$, AS A FUNCTION OF m , THE NUMBER OF NEUTRAL WORDS INTERPOSED BETWEEN THE EXPERIMENTAL WORD AND THE TEST-WORD

Circles and solid line represent means of 30 determinations; triangles and broken lines represent quartiles. Results for the control condition are shown above the letter C.

two for Condition I, in which j_a is the first contextual word and separated from j_t by two words neutral with respect to i_a . Mean values of $P_w(i)$ are indicated by the circles and solid line, quartiles by triangles and dashed lines. At the right of the graph, opposite the abscissal point labeled C, are shown the results for the control condition.

These data show that a contextual word has its greatest effect upon the association when it occurs immediately prior to the test-word. The word's effect is considerably diminished by introducing another word between it and the test-word. Interposition of two rather than one neutral word results in no appreciable further decrease in the effect of the experimental

word. In each position the contextual words have a greater effect than have the control contexts. These trends are equally clear for means and for both quartiles. Statistical analysis, the results of which are presented in Table I, bear out these conclusions.

Ideally, this experiment would consist of a very long context in which the associative relations that define the experiment were preserved. Then the dissipation of a contextual word's effect on the probability of an associative cluster could be described as a function of the number of interposed neutral words with the total number of words in the context as a parameter. For very large numbers of interposed words the effect of a contextual word should become negligible, and $P_w(i)$ should therefore approach its value in the control condition. In Fig. 1, however, the tendency

TABLE I
SIGNIFICANCE OF DIFFERENCES IN EXPERIMENT I

Values of t ($df.=29$), with corresponding p values, for differences in $P_w(i)$ when various numbers of neutral words are interposed between the experimental word and the test-word

		Number of interposed words		
		0	1	2
1	t	3.48		
	p	.01		
2	t	2.94	0.12	
	p	.01	.9	
Control	t	4.47	2.67	2.46
	p	.01	.02	.02

is for $P_w(i)$ to approach a value higher than that found in the control condition. In our opinion, this is not merely an indication of inaccuracy of the data, but results from the reinforcement of the first word of a context by an additional factor. This reinforcement can be thought of as a consequence of greater attention paid the first word of a sequence, or as a result of the fact that the first contextual word is the only one that is free of the competitive tendencies aroused by a prior word. Another possibility is that the transitional effects of the contextual words upon each other are not as negligible as assumed. The problem is amenable to straightforward experimental investigation.

Experiment 2: Density. For this experiment we take three words having strong first-order associative effects upon i (experimental words) and compare the probability of i following a context including only *one* of them with its probability following *two* of them or *three* of them.

The word sets necessary for this experiment may be specified as follows:

Condition I,	$j_a, j_1, j_2, j_t;$	
Condition II,	$j_a, j_b, j_2, j_t;$	[2]
Condition III,	$j_a, j_b, j_c, j_t;$	
Condition IV,	$j'_1, j'_2, j'_3, j_t;$	

where j_a, j_b, j_c are contextual words presumed to have strong first-order associative effects upon the same associative cluster, i , and j_1 and j_2 are words presumed to have negligible first-order effects upon i . Word-set IV refers to the control condition. An illustration can be provided by the word-sets: I, *devil, eat, basic, dark*; II, *devil, fearful, basic, dark*; III, *devil, fearful, sinister, dark*; and IV, 429, 124, 713, *dark*, where the probability of associate *hell* is to be measured. On the average, the three experimental words can be considered to have approximately equal first-order associative effects upon i , for the contextual words were assigned at random to j_a, j_b , and j_c in the various word-sets.

The probability of i following word-sets containing one, two, and three experimental words appears in Fig. 2. As in Fig. 1, circles and solid

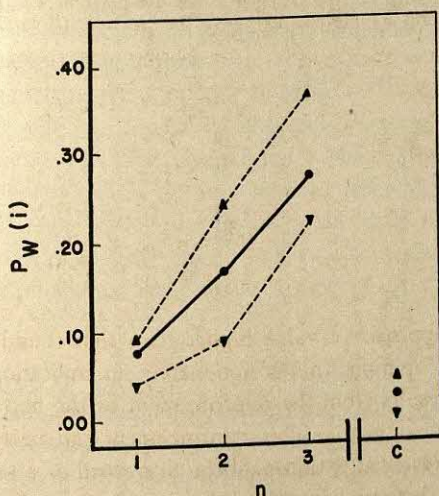


FIG. 2. THE PROBABILITY OF AN ASSOCIATIVE CLUSTER, $P_w(i)$, AS A FUNCTION OF n , THE NUMBER OF CONTEXTUAL WORDS HAVING STRONG FIRST-ORDER ASSOCIATIVE EFFECTS ON i

Circles and solid line represent means of 10 determinations; triangles and broken lines, quartiles. Results for the control condition are shown above the letter C.

lines represent means, while triangles and dashed lines represent quartiles. Each point is based on 10 measurements, one for each word-set. Results for the control condition are shown opposite the letter C. The probability of an associative cluster is seen to be an increasing function of the number

of contextual words having strong first-order associative effects upon that cluster. Statistical analysis, summarized in Table II, indicates that none of the differences in Fig. 2 would be expected to occur by random sampling as often as one time in a hundred.

In these raw data the effects of contextual words—which constitute the independent variable—are confounded with those of the test-word. These effects must be separated, for test-word and contextual words cannot be treated uniformly in view of the emphasis placed upon the former by the instructions. It has been suggested previously that the effect of a context can be measured by the *difference* between the probability of an associative cluster following an experimental word-set and either (a) its probability following the corresponding control context or (b) its first-order asso-

TABLE II
SIGNIFICANCE OF DIFFERENCES IN EXPERIMENT 2

Values of t ($df=9$), with corresponding p values, for differences in $P_w(i)$ when different numbers of words with strong first-order associative effects appear in the context.

		Number of contextual words		
		1	2	3
2	t	2.97		
	p	.02		
3	t	5.49	5.66	
	p	.01	.01	
Control	t	3.35	4.73	6.58
	p	.01	.01	.01

ciative probability following the test-word (*i.e.* its Kent-Rosanoff frequency). Either of these procedures tacitly assumes that the associative effects of context and test-word are *algebraically additive*. Since the effects of the independent variable can be measured by the present technique only if the contribution of the test-word can be extracted, an assumption of this type is indispensable.

A test of the additivity assumption can be obtained from the data presented in Fig. 2. We assume that the first-order associative effects of each member of a word-set are additive. Then the difference $\Delta P_w(i)$ between $P_w(i)$ following word-sets I, II, or III and $P_w(i)$ following the corresponding control word-set (IV) should be directly proportional to n , the number of contextual words having strong associative effects upon i . Taking the interposition variable of Experiment 1 into account, the theoretical

value of $\Delta P_w(i)$ for a context of Experiment 2 is:

$$\Delta P_w(i) = \frac{\sum_r}{r} K p_{jr}(i), \quad [3]$$

where K is a weighting factor, evaluated from the results of Experiment 1, which depends upon the number of words interposed between j_r and j_t .⁷ These weights, obtained by subtracting the probabilities of the associative clusters under control conditions of Experiment 1 from their probabilities under the corresponding experimental conditions, are 0.025, 0.028, and 0.071 for, respectively, 2, 1, and 0 interposed neutral words. Thus the following theoretical values of $\Delta P_w(i)$ for Experiment 2 are given by Equation [3]: Condition I (one experimental word), 0.025; Condition II (two experimental words), 0.053; Condition III (three experimental words), 0.124. As the design of Experiment 1 required three associative clusters for each word-set while Experiment 2 required only one, the word-sets of the latter tend to be somewhat larger. This difference in size of associative clusters can be corrected by multiplying the theoretical values by a constant which equates theoretical and experimental values for any one of the experimental conditions. Adjusted in this manner for Condition I, in which only one contextual word has a strong associative effect, the theoretical means for Experiment 2 are 0.050, 0.106, and 0.248, compared with the experimentally obtained means (data from Fig. 2) of 0.050, 0.137, and 0.245. For neither the two- nor the three-word experimental conditions does the theoretical value represent a significant departure from the experimentally obtained mean (t 's are, respectively, 1.03 and 0.08 with 9 d.f.).

This comparison is based upon the assumption that the probability of the associative cluster unaffected by contextual words can be measured by the probability of the cluster following the control context. There is some reason to believe that this method may give too high a value, however, for in a few cases the numbers of control contexts probably had an appreciable associative effect. A second comparison of theoretical and observed values for Experiment 2 was therefore made, using the relative frequency of the words of the associative clusters in the Kent-Rosanoff tables in place of the control-condition values of $P_w(i)$. Computed by this method, the theoretical means are 0.056, 0.116, and 0.228 and the empirical means are 0.056, 0.143, and 0.251. Again the t 's fail to approach statistical significance for the two- and the three-word conditions.

The assumption that associative effects are algebraically additive is thus

⁷ This assumes that the results of Experiment 1, in which the interposed words were neutral with respect to the measured associative cluster, hold also for the interposition of words having strong associative effects on it.

consistent with the present data. It can be expected to hold, however, only within the limitations of two defining conditions of these experiments: (1) that the transitional probabilities of the contextual words are negligible; and (2) that the component associative effects are not too large. The latter restriction is imposed by the logical requirement that the sum of the component associative effects be less than unity. As for the former restriction, a word-set consisting of a very familiar sequence (e.g. "*Mary had a little*") would almost certainly lead to a disproportionate number of associations of the completion or speech-habit type (e.g. *lamb*).⁸ It is also probable that the associative effects of the component words, or their weights in influencing the response, would be modified by the transitional probabilities (cf. Experiment 3 below). Even thus qualified, the assumption of additivity should be accepted only with considerable caution, since the possibility remains that even within the present defining conditions some word sets can be found that will yield results incompatible with the assumption.

Experiment 3: Frequency. The extent to which the associative response to a word-set is determined by the first-order associative effects of a particular contextual word may be expected to depend upon how familiar *S* is with that word.

In this experiment we compare the effects of two contextual words, j_a and j_b , for which $p(j_a) > p(j_b)$, where $p(j)$ is the probability of occurrence of j in a general sample of the language behavior of the population under study. The further condition is imposed that the two words have approximately equal first-order associative effects upon the associative cluster, i.e. that $p_{j_a}(i) = p_{j_b}(i)$. In lieu of the tables of first-order associative probabilities called for by this condition, a subjective approximation was attempted by selecting pairs of words that are closely synonymous, as judged by the experimenters and corroborated by a thesaurus.⁹ The following word-sets then define the experiment:

- | | | |
|----------------|--------------------------|-----|
| Condition I, | $j_1, j_2, j_a, j_c;$ | |
| Condition II, | $j_1, j_2, j_b, j_c;$ | [4] |
| Condition III, | $j'_1, j'_2, j'_3, j_c.$ | |

For an example let the synonyms *praise* and *panegyric* be the contextual words j_a and j_b and let *glory* be the association i made to the test-word *soldier*.

The mean probabilities of the 10 associative clusters obtained under each of the three conditions of this experiment are as follows: control (Condition III), 0.032; infrequent-word (Condition II), 0.035; fre-

⁸ For the meaning of these classifications cf. R. S. Woodworth, *Experimental Psychology*, 1938, 350-352.

⁹ C. O. S. Mawson, *Roger's Thesaurus of the English Language in Dictionary Form*, 1940.

quent-word (Condition I), 0.075. The differences are statistically reliable between Conditions I and II ($t = 2.64$, $p < 0.05$) and between Conditions I and III ($t = 3.46$, $p < 0.01$). The difference between Conditions II and III is not significant.

These results indicate that the frequency with which a contextual word occurs in the general language behavior of a population can be regarded as a factor weighting that word's contribution to the associative response. The insignificance of the difference between the control and infrequent-word conditions then implies that the average weight of the infrequent words in this experiment was so low that their effective contribution to the response was little greater than that of neutral items.

Next let us express the contextual word's weight in determining the associative response as a function of its frequency of occurrence. On the assumption that associative effects of contextual words are algebraically additive, the desired function is given by f in

$$\Delta P_w(i) = f[p(j)]. \quad [5]$$

As in Experiment 2, the quantity $\Delta P_w(i)$ represents the difference between the probability of associative cluster i following an experimental word-set (I or II in this experiment) and its probability following the control word-set (III). The term $p(j)$ has been defined as the probability of occurrence of word j in a general sample of the language behavior of the population under study—American college students in this case. The frequencies of words in the Lorge Magazine Count and Thorndike-Lorge Semantic Count, which correlate highly with college students' ratings of the frequency with which they use words, can be used to measure $p(j)$.¹⁰ Taken together, these counts give the number of times that a word appeared in highly varied samples of written language behavior totalling over nine million words.

Fig. 3 presents the data. The abscissa, graphed logarithmically to conserve space, gives the Thorndike-Lorge frequency of each experimental word, j_a or j_b , and the ordinate gives the difference $\Delta P_w(i)$ between the probability of associative cluster i following a word-set containing the experimental word and the probability of i in the control condition. Correlation coefficients computed for $\Delta P_w(i)$ as a function of $\log p(j)$ are significantly greater than zero ($\eta = 0.88$, $r = +0.77$); the difference between η and r is not sufficient to warrant rejection of the hypothesis that the function is rectilinear in $\log p(j)$ ($F = 1.91$; $df = 5, 13$;

¹⁰ E. L. Thorndike and I. Lorge, *The Teacher's Word Book of 30,000 Words*, 1944; Howes, *loc. cit.*

$p > 0.10$).¹¹ A rough estimate of the reliability of the measurements of $\log p(j)$ is given by the correlation between the frequencies of words in

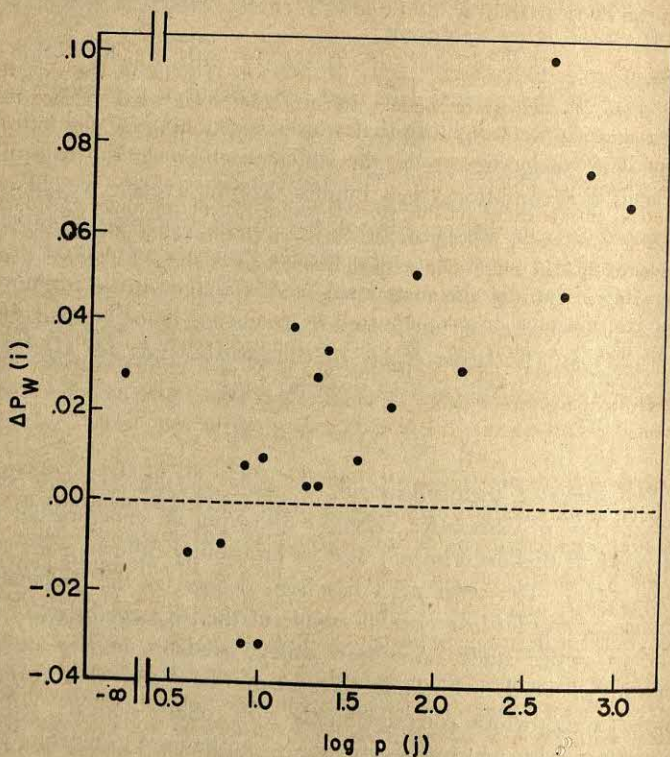


FIG. 3. THE WEIGHTED ASSOCIATIVE EFFECT OF A CONTEXTUAL WORD AS A FUNCTION OF ITS PROBABILITY OF OCCURRENCE

The abscissa shows the Thorndike-Lorge frequency of the experimental word. The ordinate gives the difference between the probability of an associative cluster following a context including the experimental word and its probability following a control context.

the Semantic Count and their frequencies in the Magazine Count ($r = +0.80$).¹² Since prediction of $\Delta P_w(i)$ from the frequency of oc-

¹¹ The small number of Ss made it necessary to estimate $P_w(i)$ for each word-set in relative-frequency units of 0.02. This is too coarse a step-interval to permit accurate estimation of the small values of $P_w(i)$ that obtain for the control word-sets. In computing $\Delta P_w(i)$ for Fig. 3, therefore, the mean probability of i over all 10 of the control word-sets has been used. This is permissible because all associative clusters and contexts for the control condition were subjected to the same selection procedures. If each value of $\Delta P_w(i)$ is recomputed using the control frequency for each individual word-set η and r are reduced to 0.71 and +0.64, respectively.

¹² The four rarest words could not be included in these calculations since their

currence of contextual words is of the same order of accuracy as prediction of word frequencies from one sample to another, a causal relation between the two variables is indicated.

Some of our Ss may not have known the meanings of all of the rare contextual words. The first-order associative effects of these words could not be expected to resemble those of their corresponding frequent contextual words. Hence the infrequent contextual words would function as neutral words, and the infrequent-word contexts would be constituted essentially like control contexts. This possibility offers an alternative interpretation of the fact that the results for the infrequent-word contexts do not differ reliably from those for the control contexts. Although it is impossible to discover directly the extent to which this factor affected the results, it seems unlikely that more than a small proportion of the Ss employed were unfamiliar with most of the infrequent words used here. The function of Fig. 3 also gives no evidence of a discontinuity such as might be expected to result from an artifact of that type. Moreover, the function approaches zero for fairly well-known words (e.g. *astrigent*, *delectable*): thus words of even greater rarity would also be expected to have zero weights, in which case it would make no difference to the experimental results whether the subjects understood the word or not.

Independence of neutral contextual words. In all of these experiments it has been assumed that the words we have selected as neutral contextual words (represented by numerical subscripts) have no appreciable effect upon the probability of the associative clusters. Precautions taken to assure satisfactory selection of these neutral words have already been explained, but it is desirable to have an empirical check.

The data of Experiment 3 can be used for this purpose. In Conditions I and II of that experiment we have two word-sets that are identical save for their third contextual word. A new associative cluster u can then be so chosen that (1) each of the first two contextual words has a strong associative effect upon u , and (2) the words in the third contextual positions are neutral with respect to u . These third contextual words in Experiment 3 are synonyms with comparable first-order associative effects upon u , but the third word of Condition I is a much more frequent word than that of Condition II. Now let us suppose that the assumption that these words are neutral with respect to cluster u is *false*. This will mean that the third word of each word-set will contribute to the probability of u . By Experiment 3, the contribution of a frequent word to the associative response must be weighted much more heavily than that of a rare word. The third contextual word of Condition I will thus increase the probability of u more than will the third contextual word of Condition II, and therefore the measured value $P_w(u)$ will, on the average, be larger for Condition I. The results, however, show a small difference in the opposite direction: the mean probability of clusters u is 0.100 following Condition I (fre-

frequencies in the Magazine and Semantic counts are not distinguished in the published Thorndike-Lorge tables; but as the scatter-diagram of the data indicates that the reliability of infrequent words is as low as, or lower than, that of frequent words, the present argument is not invalidated by their omission.

quent-word contexts) and 0.117 following Condition II (infrequent-word contexts). A difference of this size would be expected more than 3 out of 10 times by random sampling ($t = 0.98$, $df = 9$). Consequently the null hypothesis, that words selected to be neutral have in fact no effect upon an associative cluster, should not be rejected.

Discussion. The results of these experiments lend themselves to a surprisingly simple interpretation. This, however, requires a more refined definition of the *associative effect* of a stimulus-word. We consider that S is capable of emitting any one of a set of alternative word-responses $i_\alpha, i_\beta, \dots, i_v$. Each response-word we assume to have an average probability of emission $p(i)$ independent of any stimulus-word. These values may be regarded as the relative habit strengths of the words, and presumably they are sampled by tables like those published by Thorndike and Lorge.¹³ The effect of a particular stimulus-word (when S is set by the instructions of the word-association test) is then to *redistribute* these probabilities, increasing them for some words, decreasing them for others, and leaving some practically unchanged. Thus the associative effect of a stimulus-word is properly measured by the difference $p_j(i) - p(i)$. A set of such probability changes for all possible responses, $i_\alpha, \dots, i_v$, we assume to be a fixed property of the stimulus-word and the population of S s.

Consider now what happens when an S perceives two or more stimulus-words in the association-experiment. The change in probability of one response-word relative to another is a property of each stimulus-word and cannot be changed by the fact that each stimulus-word now appears as one of a sequence. Only the extent to which a stimulus-word affects the response-mechanism as a whole—its weight in Equation [3]—can vary. Thus it is *the capacity of the perceived stimulus-word to 'capture' the response-mechanism* that decreases with the number of other stimulus-words interposed between it and the moment of response (Experiment 1), and that increases in approximate proportion to the logarithm of its probability of occurrence (Experiment 3).¹⁴ Hence the high correlation between $\Delta P_w(i)$ and $p(j)$ found in Experiment 3—which would hardly be expected if more complicated interactions among the effects of con-

¹³ The correct statistic would be $\sum_j p(j)p_j(i)$, the sum of the associative probabilities for all possible stimulus-words weighted according to the probability of occurrence of each stimulus-word. A preliminary comparison of the Kent-Rosanoff and Thorndike-Lorge tables indicates that the Thorndike-Lorge frequency of a word gives a good estimate of this value except for a few special classes of words.

¹⁴ In this connection it is interesting to note that the time for which a stimulus-word must be exposed tachistoscopically in order for it to be perceived can be decreased in approximate proportion to the logarithm of its probability of occurrence (cf. Howes, *op. cit.*, esp. Ch. IV).

textual words took place—and the additivity of associative effects found in Experiment 2.

In this paper only word-sets with zero transitional probabilities have been considered. What would happen if this restriction were removed? We have already seen that, on empirical grounds, one can expect that the presence of familiar sequences in the word-sets would greatly modify the associations given. The simple model described above predicts such differences. When transitional probabilities are appreciable, the probability of occurrence of a stimulus-word, $p(j)$, depends upon the particular words that precede it. The weights of the respective contextual words in determining the responses would thus be changed greatly, yielding results for $P_w(i)$ entirely different from those calculated on the simple basis used in the present studies.

This conception of the associative process is much simpler than many views of linguistic processes would lead us to expect. It does not, for example, postulate the representational mediation-processes found necessary by one of the authors to account for many aspects of linguistic behavior, particularly those involving semantic functions.¹⁵ This simplicity, however, relates only to the manner in which certain concepts are interrelated. The concepts themselves are statistically defined and thus are recognized to be the product of complex multiple determination. Study of further variables within the present experimental design may, indeed, require a more complicated interpretation like that afforded by the mediation hypothesis.

SUMMARY

(1) The prediction of the language behavior of one population of Ss from the language behavior of a second population is formulated in statistical concepts. The basic concept is that of *associative word-probability*, defined as the probability that one person (or population of persons) will emit a word as an association following the emission of a given stimulus-word by another person. This concept is applied to the prediction of the probability of a word-association following a sequence of stimulus-words from the probability of that association following each of the component stimulus-words taken separately.

(2) Three experiments, each using 200 college students, indicate the following: (a) the effect of a given stimulus-word on an associative response is a decreasing function of the number of additional stimulus-words interposed between it and the time of response; (b) the effect of

¹⁵ C. E. Osgood, The nature and measurement of meaning, *Psychol. Bull.*, 49, 1952, 197-237.

a sequence of stimulus-words upon an associative response is an increasing function of the proportion of those stimulus-words having similar first-order associative effects on the response; and (c) the effect of a given stimulus-word on an associative response is an increasing function of the frequency of occurrence of the stimulus-word in general linguistic usage.

(3) Quantitative specification of these functions suggests certain assumptions about the way in which the effects of several different stimulus-words interact upon the same associative response. The present data are consistent with the assumption that these effects are algebraically additive.

FIGURAL AFTER-EFFECTS IN STROBOSCOPIC MOVEMENT

By BRUCE H. DEATHERAGE, University of Texas

The experiments here reported were designed to study the effects of satiation on stroboscopic movement. The dependent variable was the time between successive exposures of two stimuli which was judged to give optimal movement, and changes in this variable produced by prolonged inspection of moving and static figures were measured. The results obtained were compared with deductions based on Korte's laws and on principles derived from the examination of static after-effects in visual space.¹

EXPERIMENT 1

The first experiment was concerned with changes in the time required for optimal movement (t in Korte's laws) during the prolonged inspection of stroboscopic movement. On the assumption that apparent movement stems from a figural process in the neural locus corresponding to the line of movement, an increase in the functional distance between the end-points may be expected.² If this increase in functional distance is equivalent, as the Köhler-Wallach theory suggests, to an increase in the physical distance (s) between the points, then from Korte's statement of the relation between t and s ($t_{\text{opt}} \sim s$) it follows that t should increase in the course of prolonged inspection of movement. The study of static after-effects also gives reason to believe that the brightness of the moving figure should be reduced during continued observation, and on this basis, too, t should be expected to decrease since, according to Korte, $t_{\text{opt}} \sim 1/i$. The two effects should be separated, if possible, but to begin with it is useful to study their combined operation.

Apparatus. A Masonite screen, 8 ft. high and 8 ft. 4 in. wide, painted flat black, served as the background. There were two openings in the screen, symmetrically

* Accepted for publication May 15, 1953. This paper is adapted from a doctoral dissertation submitted to the University of Texas. The author is indebted to Professor M. E. Bitterman, who directed the research, and to Professor L. A. Jeffress, who designed the electronic equipment employed.

¹A. Korte, *Kinematoskopische Untersuchungen*, *Zsch. f. Psychol.*, 72, 1915, 193-296; W. Köhler and H. Wallach, *Figural after-effects: An investigation of visual processes*, *Proc. Amer. Phil. Soc.*, 68, 1944, 269-357.

²R. J. Christman, *Figural after-effects utilizing apparent movement as inspection figure*, this JOURNAL, 66, 1953, 66-72.

placed on either side of its center-line, which were fitted with hinged doors. Behind each opening was a box, painted flat white inside, containing a fluorescent lamp 18 in. in length. The front of each box was equipped with slots into which filters and stimulus-figures could be inserted. A fixation-point, equidistant between the stimulus-figures, consisted of a small neon lamp. The light from the lamp was filtered by a red plastic cover and could be seen through a hole ($1/32$ in. in diameter) cut in the screen.

The fluorescent lamps which lighted the stimulus-figures were powered and timed by electronic apparatus which provided a cycle with a light-dark ratio of 1:4—one light came on for $1/10$ of the cycle, followed by a dark interval of $4/10$ of the cycle, then the second light would come on for another $1/10$ cycle, and the remaining $4/10$ would be dark. The cycle was repeated, and *O* had continuous control of the rate of repetition which could be varied from less than 1 to 60 cps. The slowest rate was slow enough that succession was easily observed, while the fastest rate yielded simultaneity. Since the on-off ratio was constant, as *t* (off-time) diminished or increased, *e* (on-time) also diminished or increased. Consequently, *time in seconds per 100 cycles* (which was designated *T*) could be used as a measure of time between stimuli (the *t* in Korte's laws), the relation between the two measures being linear.

Stimulus-figures. The stimulus-figures employed were two circular apertures $1/4$ in. in diameter cut in black cardboard. The centers of these apertures were 50 in. apart. With *O* at a distance of 22 ft. from the fixation-point, the angular size of each aperture was 3 min., while the angular distance between them was $10^{\circ} 43' 30''$.

Observers. Twenty *O*s, 13 women and 7 men drawn from undergraduate classes in psychology, served in the experiment. They were divided into four subgroups of 5 *O*s each. No *O* knew the purpose of the experiment.

Task. *O* sat on a stool equipped with a chin-rest, his eyes at the same level as the fixation-point and the stimulus-circles (4 ft. 10 in. from the floor). Under these conditions the movement appeared in a line back and forth between the stimulus-circles. Beginning at a setting which gave either succession or simultaneity, *O* adjusted the repetition-rate to the point which gave him the best appearance of movement. Optimal movement was defined for *O* as occurring when the point of light (for such it appeared) seemed *clearest* during the course of movement between the end-points.

Procedure. There were three experimental sessions for each *O*. The sessions were one week apart, and each began with a 10-min. period of dark-adaptation.³ The first was a training session designed to familiarize *O* with the equipment and to acquaint him with his task. There were two criteria of performance which *O* attained during this training session. First, he learned to make a setting (judgment) for optimal movement in 20 sec. or less; most settings in the two experimental sessions were made in 10-15 sec. An interval of 20 sec. was considered to be long enough for an accurate setting and yet short enough to have but little satiating effect. Each *O* also learned to make reliable settings; the criterion set by *E* was that *T* should not fluctuate over a range greater than 10-12 sec.

³ The only light in the experimental room was furnished by three small neon lamps and a small green incandescent lamp behind the background-screen, and by another neon light at *O*'s stool.

The procedure during the two main sessions was essentially the same for all *Os*. The sessions were either 'experimental' or 'control.' Both the experimental and the control sessions consisted of 12 trials separated by 2-min. rest-intervals (except for an 8-min. rest-interval between trials 10 and 11). After each of the first 7 trials in the experimental session, the first 60 sec. of the 2-min. interval was occupied with inspection of the movement just set, while there was no satiation between settings in the control session. For half the *Os* ascending (A) and descending (D) determinations were made in the order ADDADAADADAD, while for the remaining *Os* the order was DAADADDADADA. For half the *Os* the sequence of sessions was experimental-control, and for the remaining half control-experimental. The two orders and the two sequences of conditions were balanced over the four subgroups.

Results. Table I shows the results of the experiment, in which *T* was recorded to the nearest 0.1 sec. In Fig. 1, *T* is plotted as a function of

TABLE I
MEAN TIME (*T*) PER TRIAL UNDER CONTROL AND EXPERIMENTAL CONDITIONS

Trial	(Experiment 1) Control	Experimental
1	63.1	57.4
2	63.6	62.2
3	65.3	68.5
4	64.8	69.1
5	63.0	73.3
6	64.4	70.5
7	64.5	73.2
8	61.5	74.6
9	61.4	66.4
10	66.0	66.4
11	65.4	58.3
12	64.5	60.7

paired trials (in which directional errors are balanced out) under both conditions. There is little change from trial to trial in the control session, the largest difference (4.6 sec.) occurring between trials 9 and 10. Under experimental conditions, however, *T* rises throughout the satiating period (a mean difference between Trials 1 and 8 of 17.2 sec.), and then decreases with the termination of the satiating procedure on Trial 8. The difference between the settings made in experimental and control sessions on Trials 5-6 were significant beyond the 5-% level (*t* for correlated data = 2.75 with 19 *df*) as was the difference between settings made on Trials 7-8 (*t* = 2.85). The difference between settings made in the experimental session on Trials 1-2 and 7-8 was significant beyond the 1-% level (*t* = 5.54).

The data here presented clearly suggest that prolonged inspection of

stroboscopic movement gives rise to a significant change in the time between stimuli required for 'best' movement to be perceived; the prediction based on the Köhler-Wallach data and Korte's laws thus seems to be fulfilled. Apparently, functional changes in intensity and distance are equivalent to corresponding physical changes with respect to their effect on the time required for optimal movement.

In a preliminary report on the effect of satiation on stroboscopic movement produced with an interocular apparatus, Deatherage and Bitterman described results which are contradicted by the present experiment.⁴ They noted that after satiation "succession is reported at a rate of alternation

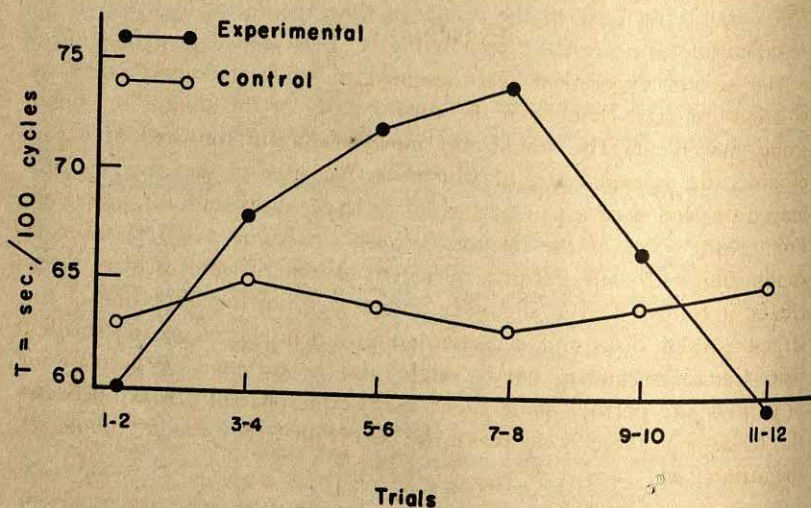


FIG. 1. CHANGES IN TIME FOR OPTIMAL MOVEMENT UNDER EXPERIMENTAL AND CONTROL CONDITIONS IN EXPERIMENT I

previously set for optimal movement. In some *Os* movement can be reinstated by increasing the rate of alternation, while for others an increase in rate yields simultaneity without intervening movement." In the light of the present experiment, it is clear that the preliminary report was not simply wrong, but merely incomplete. The earlier assumption was that if succession were observed, then obviously the rate of alternation was too slow (*T* too long), and it was quite natural for *Os* to endeavor to reinstate movement by increasing the rate. Only some *Os*, however, reported

⁴ B. H. Deatherage and M. E. Bitterman, The effect of satiation on stroboscopic movement, this JOURNAL, 64, 1951, 108-109.

reinstatement of movement—a report which now appears to be an error of expectation—while others reported almost immediate simultaneity. With opportunity given to search for best movement up and down the alternation-rate scale, the *O*s of the present experiment found movement at a slower rate. Herein exists a difference between physical changes and satiation-produced functional changes in distance and intensity. While a physical increase in distance or decrease in intensity could be expected to produce simultaneity, the corresponding satiation changes produce reports of succession, although the change in *T* required to reinstate optimal movement is in the same direction (decrease) in both cases. No explanation of this discrepancy can be offered at present.

EXPERIMENT 2

The second experiment was designed to study the effect on *T* of satiation-produced changes in the apparent size of the stimuli for stroboscopic movement. The stimuli for movement again consisted of a pair of disks, the apparent size of which was increased or decreased by prolonged inspection of a pair of smaller or larger disks statically presented. From Korte's law of the relation between *t* and *i* ($t_{\text{opt}} \sim 1/i$) taken in conjunction with empirical principles derived from the study of static after-effects, it follows that *T* should be increased following inspection of the largest pair of disks and decreased following inspection of the smallest pair. Certain secondary effects might also be expected—a decrease in brightness and perhaps some slight increase in apparent distance between the disks—but there is no reason to believe that they should counteract the main effect.

Apparatus. The apparatus used was essentially that of the first experiment, with a few changes made in the relationship between *O* and the stimulus-field. To facilitate the perception of stroboscopic movement, which occurs more readily in peripheral vision, the fixation-point was moved from the position occupied in the first experiment to a point 14 in. higher,⁵ but still in the centerline of the screen. The distance between *O* and the screen was reduced to 215 in. (17 ft. 10 in.), and to preserve the original angular separation the distance between the centers of the circular apertures employed was reduced to 40.75 in.

Stimulus-figures. The stimulus-figures were three in number: (1) a pair of small circular disks, each 0.5 in. in radius; (2) a pair of large disks, each 2.0 in. in

⁵ Different areas of the retina may show different basal speeds for stroboscopic movement—a problem still to be investigated. This may be the reason for the approximately 20-sec. difference between *T*s for control trials in the two experiments. A tendency (not critical) was also noted in the first experiment for the movement to 'break' in the region of the fixation point when that point lay in the path of the movement.

radius; and (3) a test-figure of two intermediate disks, each 1.25 in. in radius.⁶ The disks were presented as in Experiment 1, appearing as light figures on a dark ground. The left-hand member of each pair of disks was always centered at the same point, as were the right-hand members of each pair, the two points being $10^{\circ} 43' 30''$ apart. As in the previous experiment, the on-off ratio was 1:4.

Observers. The Os were 10 young men who were selected from undergraduate courses in psychology. None knew the purpose of the experiment, and none had participated in Experiment 1.

Procedure. There were three experimental sessions for each O, approximately a week apart. The first session, as in Experiment 1, was utilized to acquaint O with the apparatus, and to have him learn to meet the same two levels of performance demanded in the earlier experiment. In addition, there were two experimental sessions (Conditions 1 and 2).

In *Condition 1* the satiation-figure consisted of the small (0.5-in.) disks, while in *Condition 2* the large (2.0-in.) disks served. In both conditions 12 trials were

TABLE II
MEAN TIME (T) PER TRIAL UNDER TWO CONDITIONS OF SATIATION

Trial	(Experiment 2)	
	Condition 1	Condition 2
1	45.8	45.3
2	43.2	46.0
3	43.8	48.1
4	47.1	47.9
5	47.6	45.8
6	50.6	49.4
7	55.8	47.1
8	53.5	49.2
9	58.0	48.3
10	57.9	47.8
11	45.7	50.6
12	44.8	48.0

given, with ascending and descending determinations balanced by means of a double-alternation order. Trials 1-4 and 11-12 were 'control' trials, and Trials 5-10 were 'experimental' trials. Each experimental trial was preceded by a 60-sec. period of satiation (with the largest or the smallest pair of disks) under conditions of clearcut simultaneity (the repetition-rate set at 60 cps). On each of the trials disks of intermediate size (1.25 in.) were used, and O was asked to set the repetition-rate to give the best appearance of movement. On experimental trials an interval of 20-30 sec. separated the satiation period from O's setting. A 2-min. rest followed each trial, except that there was a 5-min. rest between Trials 10 and 11, and a 3-min. rest between Trials 11 and 12. Half the Os worked under *Condition 1* in the second session and under *Condition 2* in the third, while the others encountered the two conditions in reverse order.

Results. The results of the second experiment are summarized in Table II, and, in Fig. 2, *T* is plotted as a function of paired trials. The curves for

⁶ At 215 in., the difference in radius from each disk to the next was 12' of arc, a value selected (on the basis of the Köhler-Wallach data) to maximize the after-effect.

both conditions (Conditions 1 and 2) remain essentially unchanged for control trials (Trials 1-4). With satiation, the curves diverge markedly, *but not in the expected direction*. On the assumption that the disks used for testing would be functionally increased in size as a result of satiation with the smaller pair (Condition 1), a decrease in T was expected. Instead, a marked *increase* appeared during satiation, followed by a decrease on Trials 11 and 12 after periods of rest. Satiation with the larger disks (Condition 2) was expected to produce a functional decrease in the size and brightness of the disks used for testing, and consequently an increase

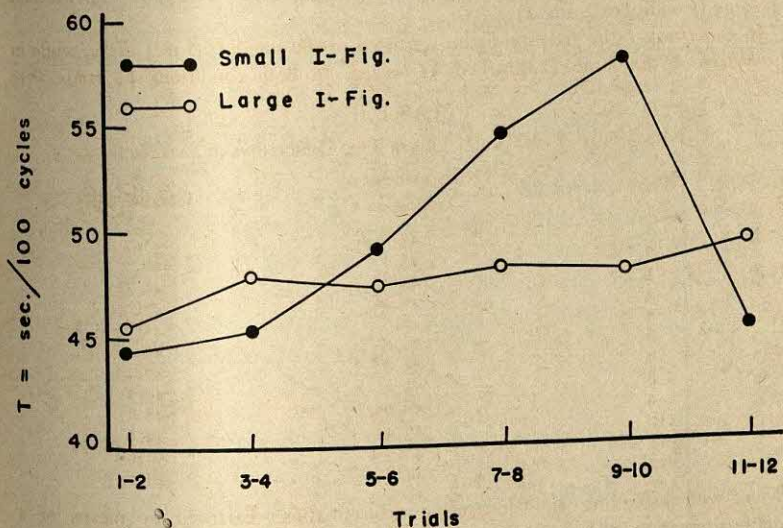


FIG. 2. CHANGES IN TIME FOR OPTIMAL MOVEMENT UNDER TWO CONDITIONS OF SATIATION IN EXPERIMENT 2

in T . It can be seen, however, that the values for T remain essentially *unchanged* throughout the 12 trials of Condition 2. A test of the significance of the difference between the two experimental conditions was performed, using the data of Trials 9-10. Each O 's settings for these trials were averaged to provide a mean setting for each condition. The difference between the two distributions of scores was significant beyond the 1-% level (t for correlated data = 4.79 with 9 df).

Reports made by O s during the satiation trials made abundantly clear the success of the satiating procedure. Many O s spontaneously remarked upon the phenomenal changes in the size of the test-disks after inspection-periods; the reports were *always* that size had changed in the expected

direction. Nevertheless, while predicted changes in size occurred, the predicted changes in T did not. This discrepancy can be accounted for in one of two ways: either Korte's statement of the relation between i and t is incorrect; or changes in the physical size of the stimuli for movement are not functionally equivalent to changes in phenomenal size produced by satiation. In order to check the validity of Korte's law, determinations of T for optimal movement were made with stimuli which differed in physical size.

The relation between T and physical size. Five Os, all men, were used. Each was dark-adapted for 10 min., and acquainted with the apparatus and the task, following which six determinations were made—one ascending and one descending setting for each of the three pairs of disks employed in Experiment 2.

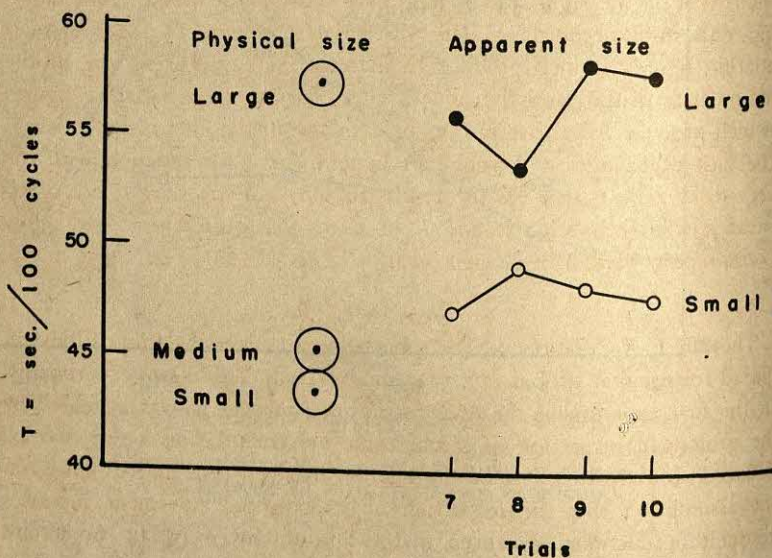


FIG. 3. TIME FOR OPTIMAL MOVEMENT AS AFFECTED BY DIFFERENCES IN PHYSICAL AND IN APPARENT SIZE (EXPERIMENT 2)

The determinations were made in random order. The mean T s for the small, medium, and large disks were 43.4, 45.3 and 57.3, respectively. The difference between the first two falls far short of statistical significance. There is, however, a clear difference between T -values for the medium and large disks. By a simple binomial test the difference was significant beyond the 4-% level, the probability being 1/32 that the five differences all will be in the same direction. A t -test for correlated data showed the difference to be significant well beyond the 5-% level ($t = 4.202$ with 4 df).

The results indicate that the relation between t and i postulated by Korte does not hold, at least under the conditions of this experiment. Korte's method of increasing intensity, since he was dealing with drawn figures, was to increase the width of a line or to put short perpendiculars across its ends; the discrepancy between Korte's law and the data here obtained may be due to the difference in the method of varying the size, and thus the intensity. Nevertheless, when the effect on T of changing physical size (as here obtained) is compared with the effect on T of change in apparent size by satiation, a striking consistency appears. In Fig. 3, the correspondence of the two sets of results is demonstrated. The three encircled points on the left side of the figure show the average T required for optimal movement with the disks of different physical size. The curves on the right side of the figure show T for the last four satiating trials in Experiment 2. The top curve represents T 's for the small satiating figures (which produce an increase in the apparent size of the intermediate disks), and the bottom curve for the large satiating figures (which produce a decrease in the apparent size of the intermediate disks). The lack of change in T found in Experiment 2 for satiation with the large disks appears due (as the results for physical size show) to a curvilinear relation between size and T ; it seems that decreases in size below a certain point yield little change (if any) in T .

SUMMARY

The effect of satiation on the time between successive stimulations required for optimal stroboscopic movement was studied in two experiments. In the first experiment, the time for optimal movement was increased by the prolonged inspection of stroboscopic movement. This result may be deduced from Korte's law of the relation between time and distance, on the assumption that the inspection of stroboscopic movement increases resistance (functional distance) in the line of movement. In the second experiment, the effect on the time for optimal movement of satiation-produced changes in the size of the stimuli for movement was measured. An increase in apparent size increased the time and a decrease in apparent size had no significant effect, neither result being in accord with predictions made from Korte's law of the relation between intensity and time. Subsequent measurements of the relation between time and physical size contradicted Korte's law, however, and provided results directly comparable to those of the satiation experiment. In both experiments, then, stroboscopic movement was found to be similarly affected by physical and satiation-induced functional changes in selected parameters.

PSYCHOMOTOR PERFORMANCE AS A FUNCTION OF INITIAL LEVEL OF ABILITY

By BRADLEY REYNOLDS and JACK A. ADAMS, Perceptual and Motor Skills
Research Laboratory, Human Resources Research Center

The fact that little attention has been paid to individual differences as independent variables in learning studies is probably attributable to the interest of workers in the field of learning in the discovery of general laws and to the implicit assumption that such general laws based upon sample statistics will hold throughout the range of learning ability. There is little experimental basis, however, for this assumption, the validity of which the present study was designed to test.

There are a number of practical implications to be derived from information on learning and individual differences. Mass training problems invariably involve some personnel whose initial level of ability is low, and it becomes desirable in these instances to know whether any interaction between learning variables and individual differences exists. Further information with respect to individual differences in learning could illuminate Reynolds' finding that (a) correlations between tests and criterion-measures shrink as training on the criterion-skill progresses, and that (b) correlations between initial and subsequent measures of performance decrease as training progresses.¹ Woodrow also has carried out investigations in the area of individual differences and learning.² He has been concerned with such problems as the isolation of a general ability to learn, the effect of practice upon inter-correlations of tests, and factor analyses of learning data. Kientzle, a student of Woodrow's, examined the effect of practice and distribution of practice upon between-S variability and intertrial correlation.³ British investigators have shown some concern about the

* Accepted for publication March 23, 1953. The experimental work was performed as part of the United States Air Force Human Resources Research and Development Program. The opinions or conclusions contained in this report are those of the authors. They are not to be construed as endorsed by or reflecting the views of the Department of the Air Force.

¹Bradley Reynolds, The effect of learning on the predictability of psychomotor performance, *J. Exper. Psychol.*, 44, 1952, 189-198.

²Herbert Woodrow, The relation between abilities and improvement with practice, *J. Educ. Psychol.*, 29, 1938, 215-230; The effect of practice on groups of different initial ability, *J. Educ. Psychol.*, 29, 1938, 268-278; The effect of practice on test intercorrelations, *J. Educ. Psychol.*, 29, 1938, 561-572; Interrelations of measures of learning, *J. Psychol.*, 10, 1940, 49-74; The ability to learn, *Psychol. Rev.*, 53, 1946, 147-158.

³M. J. Kientzle, Properties of learning curves under varied distributions of practice, *J. Exper. Psychol.*, 36, 1946, 187-211; Ability patterns under distributed practice, *ibid.* 39, 1949, 532-537.

problems of learning and individual differences. Farmer, in a study which is probably closely related to the present one, found that, in the case of psychomotor skills learned in industrial situations, Ss differentiated on the basis of initial level of ability showed parallel acquisition curves.⁴

In the study which is reported here, one group of Ss was trained with massed and a second group with distributed practice. Each group was divided into subgroups on the basis of initial levels of ability in the task, and learning curves obtained for the various subgroups were examined for systematic differences in performance.

METHOD

Apparatus. The apparatus employed was the Rotary Pursuit Test which has been described elsewhere.⁵ Twelve copies of this test were used. The training trials and interpolated rest-periods were cycled automatically. The disks turned and a buzzer sounded for 1.5 sec. before the clocks could be activated. Scores were recorded in units of 0.001 min. on Standard Electric timers.

Subjects. The Ss were 960 basic airmen trainees drawn from the population available at Lackland AFB, Texas. Subgroups of 12 Ss were assigned to either a massed or a distributed practice group. The massed group had a mean age of 20.13 yr. with an *SD* of 1.47 and the distributed group had a mean age of 20.23 yr. with an *SD* of 1.78. The mean years of education for the massed group was 11.06 with an *SD* of 1.89. The distributed group had 11.27 mean years of schooling with an *SD* of 1.81.

Experimental procedure. Ss were trained with massed or distributed trials through several training sessions. It was necessary to give practice within a half day and, because of time-limitations, to train distributed groups in the forenoon and massed groups in the afternoon. The duration of each trial was 20 sec. for both groups. The intertrial rest was 5 sec. for the massed group and 60 sec. for the distributed group. The massed group was given four blocks of 20 trials each and the distributed group had three blocks of 20 trials each. There was a 30-min. rest between blocks.

When the collection of data had been completed, Ss within each experimental group were stratified into deciles on the basis of their total score for the first five trials.

RESULTS

For each group, Table I gives the distribution of scores on which the stratification into ability-subgroups was made. Skewness is to be observed in both distributions and is more marked in the case of massed than distributed practice.

Mean scores for the massed and distributed groups on all training trials are plotted in Fig. 1. It will be noted that the score of the massed

⁴Eric Farmer, Parallelism in curves of motor performance, *Brit. J. Psychol.*, 42, 1927, 335-342.

⁵A. W. Melton (Ed.), *Apparatus Tests*, AAF Aviation Psychology Program Research Report No. 4, 1947, 329-399.

group was somewhat above that of the distributed group on Trial 1. This difference, which approaches significance at the 5-% level ($t = 1.94$

TABLE I
FREQUENCY DISTRIBUTIONS OF TOTAL SCORE ON FIRST FIVE TRIALS FOR MASSED
AND DISTRIBUTED GROUPS

Interval (m. min.)	Frequency	
	massed	distributed
0- 99	52	15
100- 199	101	48
200- 299	104	78
300- 399	85	99
400- 499	75	105
500- 599	31	56
600- 699	19	49
700- 799	6	18
800- 899	5	8
900- 999	2	3
1000-1099		1

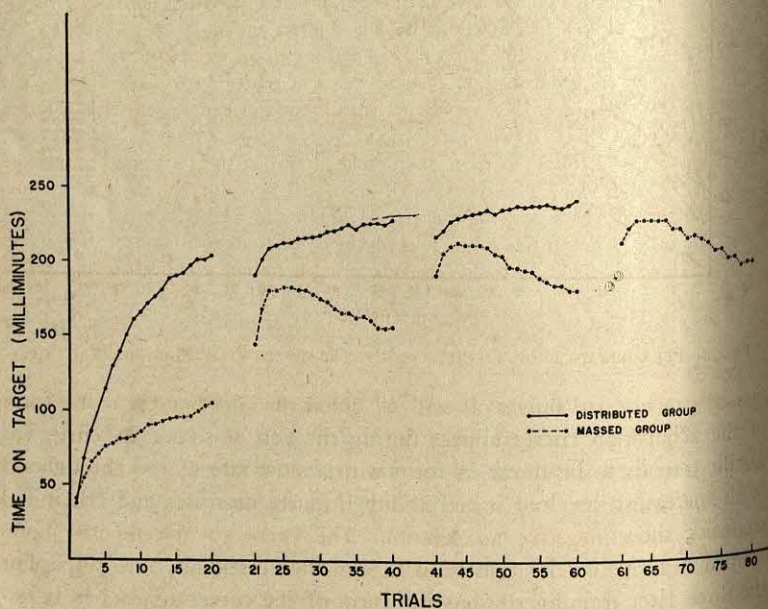


FIG. 1. PERFORMANCE OF MASSED AND DISTRIBUTED GROUPS.

for 958 df), may be due in part to sampling and in part to the difference in time of testing. The observed difference, however, does not affect the

principal analyses since the main comparisons are within-groups. The curves obtained under the two conditions of distribution are similar to those usually observed with massed and distributed trials on the Rotary Pursuit Test. There is, in the case of distributed practice, a between-session loss in proficiency. The massed group shows reminiscence and an initial post-rest warm-up segment followed by the characteristic downward trend in the later trials of each session. The marked difference observed between the groups at the end of each session is dissipated to a great extent in the early trials of the following session.

In Fig. 2, the acquisition curves for the ten decile groups trained with

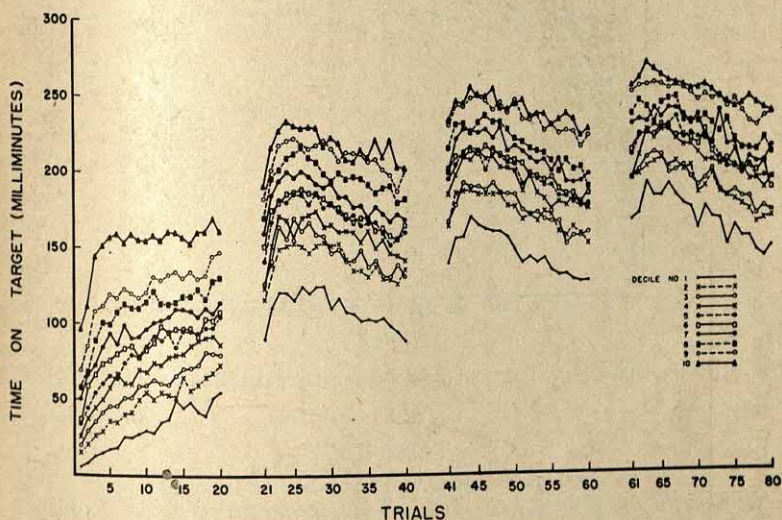


FIG. 2. PERFORMANCE OF DECILE-GROUPS TRAINED WITH MASSED PRACTICE.

massed practice are shown. It will be noted that differences in the form of the acquisition curves appear during the first session. The curve for Decile 1 tends to be linear in form with a slow rate of rise throughout. With increasing level of initial ability linearity decreases and the initial segments show negative acceleration. The curve for Decile 10 shows marked negative acceleration initially, and then essentially zero slope. For the three later training sessions the form of the curves appears to be essentially the same. All show reminiscence following rest, a period of rapid improvement in the initial post-rest trials, and a decremental segment in the later trials of each training session. In general, it would appear that for the final three training sessions the combined curve for the massed

group in Fig. 1 is fairly representative of the curve for each of the deciles in Fig. 2.

In Fig. 3, performance curves for the decile groups trained with dis-

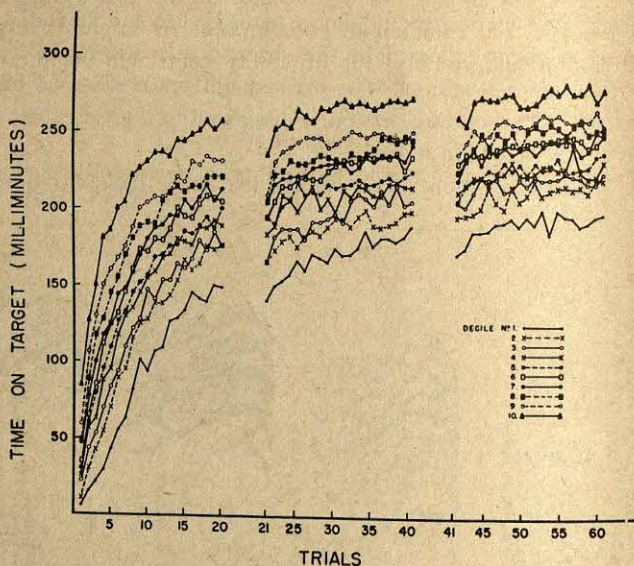


FIG. 3. PERFORMANCE OF DECILE-GROUPS TRAINED WITH DISTRIBUTED PRACTICE.

TABLE II
ANALYSIS OF VARIANCE OF SCORES FOR TRIAL 60

Source of variance	df	Sum of squares	F*	P
Columns (distribution)	1	.844	422.00	< .01
Rows (ability level)	9	.586	32.55	< .01
Within cells	940	1.848		
Remainder	9	.025	1.40	> .05
Total	959	3.303		

* Within cells mean square used as error term.

tributed practice are shown. As with massed practice there are differences in slope during the first session which are related to initial level of ability. The slopes become greater with increasing level of ability. Again, as in the case of massed practice, there would appear to be a family of curves differing in slope-constants. The curves during the second and third sessions are essentially the same. All show the same trend and all but one display

between-session loss (Decile 3). With the exception of the first session, the combined curve for the distributed group in Fig. 1 is quite representative of the curve for each of the deciles in Fig. 3.

Table II presents an analysis of variance of the scores on Trial 60 which is the last trial common to both groups. As might be expected, both distribution and initial level of ability contribute significantly to

TABLE III
FREQUENCY-DISTRIBUTIONS OF SCORES FOR TRIAL 80 FOR MASSED GROUPS
Decile

Interval (m. min.)	1	2	3	4	5	6	7	8	9	10
0-19										
20-39	1	1								
40-59			2							
60-79	4	1	1	1						
80-99	7	3	4	1	1	1	1			
100-119	3	4	3		1	3			1	
120-139	5	4	2	3	3	2	3			2
140-159	7	8	6	6	7	6	4	3		
160-179	7	7	4	8	7	6	1	8	1	1
180-199	5	7	11	6	6	11	8	8	8	3
200-219	5	7	7	8	7	12	9	13	8	6
220-239	3	4	3	8	10	3	5	8	10	13
240-259	1		4	6	4	1	11	5	12	12
260-279		1	1		1	2	4	2	4	6
280-299		1		1	1	1	1	1	4	4
Mean	147	166	170	192	192	185	212	203	231	235
SD	51	51	53	42	43	42	46	33	34	37

variability. Of greatest interest is the fact that interaction between distribution and initial level of ability does not produce a significant F . In evaluating the analysis, it must be remembered that S s were not randomly assigned to the two conditions. An apparent affect of this non-randomness was a difference on Trial 1 in favor of the massed group. By Trial 60, however, the distribution variable has separated the two groups markedly in favor of the distributed group as indicated by the highly significant F for distribution.

Tables III and IV present frequency-distributions of final-trial scores for massed and distributed practice respectively. It is noteworthy that lower-decile distributions tend to overlap upper decile distributions at advanced levels of training. Increased training seems to produce greater differentiation among S s whose initial performance is very low, as can be observed most markedly when Deciles 1 and 10 are compared in either

Table III or IV. These two groups, which are separated by 80% of the Ss in the first five trials, show a high percentage of overlap on the final trial under either massed or distributed practice, and the level of variability for lower deciles is greater than for higher deciles under both massed and distributed practice. To evaluate the possibility is that these differences in variability are attributable to differences in mean level, a trial was determined for each decile where means were equal or nearly equal. The mean and *SD* for each of these trials is given in Table V, which

TABLE IV
FREQUENCY-DISTRIBUTIONS OF SCORES FOR TRIAL 60 FOR DISTRIBUTED GROUPS

Interval (m. min.)	Decile									
	1	2	3	4	5	6	7	8	9	10
0-19										
20-39										
40-59	1				1					
60-79		2								
80-99	1		1		1					
100-119	1	1		1						
120-139	2					1	2		1	
140-159	3	3		1	2		1	1		
160-179	8	1	6	4		1				
180-199	7	5	6	8	3	6	2	2	2	
200-219	8	7	10	6	5	1	2	3	4	1
220-239	7	10	5	3	7	8	7	5	7	2
240-259	9	10	9	11	13	8	14	15	11	9
260-279	1	6	6	6	9	13	15	13	12	13
280-299		3	3	8	7	8	4	6	8	12
300-319			2			2	1	3	3	11
Mean	197	219	224	230	236	248	245	253	254	277
SD	49	50	43	44	49	38	37	31	36	25

permits some determination of the effect of initial level of performance upon variability when all groups have been equated for level of proficiency. It will be observed that there is a general tendency for lower deciles to show greater variability.

For the massed deciles, gain over each 30-min. rest-period, warm-up decrement at the beginning of Sessions 2-4, and loss within Sessions 2-4, were evaluated by simple analysis of variance to test the hypothesis that there were no differences among ability groups with respect to these measures. Gain during rest was taken as the difference between the final pre-rest trial and the initial post-rest trial. Warm-up decrement was considered as the steep initial slope of the post-rest curve and was measured as the difference between the initial trial of a session and the peak-trial of that session. The amount of loss within a session was taken as the

difference between the peak-trial and the final trial of the session. None of the F -ratios obtained were significant at the 5-% level, indicating no evidence of interaction between gain, warm-up, or loss within a session, and initial level of ability.

Warm-up decrement and gain within-session for Sessions 2 and 3 for the distributed deciles also were evaluated by simple analysis of variance. Since warm-up is manifested as an actual loss under distributed practice, it is measured as the difference between the final pre-rest trial and initial post-rest trial. Gain within a session is measured as the difference between the final trial of a session and the final trial of the preceding session. As with the massed groups, all F -ratios lacked significance at the 5-% level and gave no supporting evidence for interaction between these performance characteristics and initial level of ability.

Product-moment correlations for all S s in each of the two major groups were

TABLE V

SD ON TRIALS WITH APPROXIMATELY EQUAL MEANS FOR MASSED AND DISTRIBUTED GROUPS
(Values in m. min.)

		Decile									
		1	2	3	4	5	6	7	8	9	10
Massed											
Trial No.		63	44	44	79	51	26	29	37	39	21
Mean		189	186	188	186	189	187	188	188	185	189
SD		54	42	52	43	42	40	52	39	34	37
Distributed											
Trial No.		45	36	25	18	19	13	11	9	8	5
Mean		186	188	187	187	186	188	188	188	187	186
SD		52	50	40	41	48	37	34	35	33	28

computed, using total score on Trials 1-5 as the X -variable and total score on Trials 6-10, 56-60, and 76-80 (massed group) as Y -variables. The correlation between Trials 1-5 and 6-10 was 0.85 for distributed practice and 0.89 for massed practice. The r between Trials 1-5 and 56-60 was 0.57 for distributed practice and 0.61 for massed. The r between Trials 1-5 and 76-80 was 0.57 for the massed group. The higher correlation between total score on initial five trials and adjacent five trials, as compared with that between initial five trials and final five trials, corroborates the findings of other investigators.

DISCUSSION

With the exception of slope-characteristics of first-session curves no evidence has been found for the interaction of ability-level and learning variables. The data are consistent with those of Farmer,⁶ and support the assumption of Hull that individual differences affect the constants of

⁶ *Op. cit.*, 335-342.

a behavioral equation but not the general mathematical form.⁷ In the case of the first-session curves, the general form is apparently not affected by initial level of ability. The main characteristic which does seem to be affected is slope of the initial segment which increases with increasing level of ability.

One interesting aspect of the acquisition curves bears consideration. It was noted that the curve of massed Decile 10 displayed in the first session a very steep slope over the first few trials and then leveled off to a zero slope. This curve has an interesting similarity to certain post-rest performance curves obtained under conditions of massed practice on the Rotary Pursuit Test. If the warm-up period can be regarded as a period of recruitment of previously acquired responses, then Ss in massed Decile 10 would appear to be activating a pool of previously acquired relevant responses, carried over, perhaps, from psychomotor tasks encountered in everyday situations.

It was noted in the introductory section that the present results would furnish information relevant to problems in the prediction of advanced levels of proficiency. Two features of commonly observed results were mentioned: (1) the tendency for correlations between selection-tests and criterion-measures to shrink as training on the criterion-skills progresses; and (2) the tendency (also found here) for correlations between performance on initial and later trials to decrease as training progresses. The difference in variability between higher and lower deciles is relevant to both of these problems. It was found that distributions of scores for final trials were more variable for lower deciles than for higher deciles. It was also noted (in Table V) that the distribution of scores for deciles on trials having common means show greater variability for the lower than for the higher deciles. The tendency for variability to be greater for lower deciles does not, therefore seem to be a function of correlation between variability and the mean. Consequently, it cannot be claimed that the greater variability shown by lower deciles arises because higher deciles are approaching the task-asymptote. Further, Tables III and IV revealed a considerable overlap in the frequency distributions of final-trial scores. These findings suggest that attenuation in correlation with training arises because some Ss with initially low scores show relatively rapid rates of improvement and, in the final stages of practice, perform at a level comparable to Ss whose initial level of proficiency was much higher. One possible explanation of this phenomenon is that some Ss are initially poor performers because of lack of opportunity in the past to acquire habits which generalize to the skill being learned. When given an opportunity to practice a specific skill, they can, with sufficient training, attain the level of Ss who have had the benefit of relevant prior experience.

An interesting question to consider is whether all Ss under either massed or distributed practice will eventually approach a common asymptote. Even though some Ss change status with training, an examination of Figs. 2 and 3 shows that the mean curves are essentially parallel in the final session, with little tendency to converge. It appears that all Ss would not attain the task asymptote but rather would attain an asymptote commensurate with their own ability. The possibility exists, however,

⁷ C. L. Hull, The place of innate individual and specie differences in a natural-science theory of behavior, *Psychol. Rev.*, 52, 1945, 55-60.

that the correlation between initial and final measures of performance will continue to decrease as training progresses. If, with extensive practice, this correlation approaches zero then differentiation between *Ss* of low and high ability ceases to exist. If, however, the correlation stabilizes at some intermediate positive value, then some differentiation between low and high ability persists at these advanced levels. Additional experimentation involving prolonged practice is needed to answer these questions.

SUMMARY

Two groups of *Ss* were trained on the Rotary Pursuit Test, one with massed and one with distributed practice. Each group was subdivided into deciles on the basis of initial performance, and the learning curves of the various subgroups were compared. The form of the curves did not, for the most part, vary significantly with level of ability. The results of the experiment support the assumption that general laws of learning based on sample statistics hold throughout the range of learning ability.

FIGURAL AFTER-EFFECTS IN AUDITORY SPACE

By JOHN KRAUSKOPF, University of Texas

The Köhler-Wallach theory of figural after-effects represents the first translation of the principle of isomorphism into concrete neurophysiological terms.¹ Although still in a rudimentary stage, it seems able to encompass much of the data on the perception of contours. The distortions of visual space following prolonged fixation of visual objects, which were first reported by Gibson² and further investigated by Köhler and Wallach, provide the primary data for the theory.³ The theory has been elaborated only for two-dimensional vision, but if the assumption is made that the Köhler-Wallach model is a mechanism associated with place-representation in the cortex, similar shifts should be expected in other perceptual dimensions. The existence of figural after-effects in the visual third dimension reported by Köhler and Emery⁴ may be taken as evidence for the existence of a mechanism underlying the perception of depth similar to that proposed by Boring.⁵ The discovery by Deutsch of after-effects in auditory pitch,⁶ and the kinesthetic after-effects reported by Gib-

* Accepted for publication April 2, 1953. This paper is adapted from a dissertation submitted to the University of Texas in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The research was directed by Professor M. E. Bitterman.

¹ Wolfgang Köhler and Hans Wallach, Figural after-effects: an investigation of visual processes, *Proc. Amer. Phil. Soc.*, 88, 1944, 269-357.

² J. J. Gibson, Adaptation, after-effect and contrast in the perception of curved lines, *J. Exper. Psychol.*, 16, 1933, 1-31.

³ Recently Osgood and Heyer, elaborating on the work of Marshall and Talbot, have developed a statistical theory of figural after-effects which is based on a more orthodox conception of neurophysiology than that of Köhler and Wallach (C. E. Osgood and A. W. Heyer, Jr., A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1952, 98-118; W. H. Marshall and S. A. Talbot, Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity, in H. Klüver (ed.), *Visual Mechanisms, Biol. Sympos.*, 7, 1942, 117-164). The model of Osgood and Heyer deals effectively with much the same data as that of Köhler and Wallach, and no contradictory predictions from the two theories have as yet been made.

⁴ Köhler and D. A. Emery, Figural after-effects in the third dimension of visual space, this JOURNAL, 60, 1947, 159-201.

⁵ E. G. Boring, *The Physical Dimensions of Consciousness*, 1933, 118-119.

⁶ J. A. Deutsch, A preliminary report on a new auditory after-effect, *Quar. J. Exper. Psychol.*, 3, 1951, 43-46.

son⁷ and by Köhler and Dinnerstein,⁸ support the view that figural after-effects reflect a general characteristic of place-systems.

The case in auditory space is less clear, since no anatomical projection area has been as yet reported. Jeffress, however, has made an argument for a place theory of auditory space,⁹ proposing a neural model for the conversion of difference in time of arrival at the two ears into place. If we accept his argument and the notion that figural after-effects are a property of place-systems, it is reasonable to look for such effects in auditory space perception.

Flugel has done the pioneer research in this area.¹⁰ In experiments designed to study 'fatigue' in the auditory system he found that monotic stimulation resulted in a shift of the dichotically determined apparent medial plane. Flugel's *O*s were asked to adjust the phase-difference of a dichotic tone until it appeared medial. Following monotic stimulation, the medial plane determinations were repeated, and it was found that the phase-difference was adjusted by the *O*s so as to simulate a sound coming from the side of the 'fatiguing' tone. Flugel concluded from this evidence that monotic stimulation fatigued the peripheral auditory apparatus.

Bartlett and Mark contested Flugel's hypothesis and attempted to demonstrate the central nature of the phenomenon.¹¹ They repeated his experiments, using both monotic and dichotic fatiguing stimuli, with the dichotic stimulus adjusted in phase so that it appeared to be coming directly from one side. In these experiments Flugel's results with monotic stimulation were corroborated, but similar results with dichotic 'fatigue' also were obtained. Both sets of results were attributed to a central effect.

More recently Jones and Bressler performed similar experiments which were specifically designed to test the Köhler-Wallach conception.¹² Their test was the same as that employed in the earlier studies (*i.e.* measurements of shifts in the apparent medial plane as a function of prolonged stimulation) but they were particularly interested in demonstrating the existence of a 'distance-paradox' similar to that obtained by Köhler and Wallach in vision. For this reason, they used two fatiguing tones, one 90° and the other 60° to one side of the medial plane. They found shifts of the medial plane in the direction of the fatiguing stimuli following satiation at both points, with a greater shift at 60°. These data were taken as evidence of a distance-paradox in auditory space analogous to that found in vision.

While these results suggest that distortions can be induced in auditory

⁷ Gibson, *op. cit.*, 1-31.

⁸ Köhler and Dorothy Dinnerstein, Figural after-effects in kinesthesia, in *Miscellanea Psychologica Albert Michotte*, 1947, 196-220.

⁹ L. A. Jeffress, A place theory of auditory localization, *J. Comp. Physiol. Psychol.*, 41, 1948, 35-39.

¹⁰ J. C. Flugel, On local fatigue in the auditory system, *Brit. J. Psychol.*, 11, 1920-1921, 105-134.

¹¹ F. C. Bartlett and H. Mark, A note on local fatigue in the auditory system, *Brit. J. Psychol.*, 13, 1922, 215-218.

¹² F. N. Jones and E. Bressler, Displacement after-effect in auditory localization, *Amer. Psychol.*, 4, 1949, 389.

space, more evidence is required for the establishment of an analogy to figural after-effects in vision. The experiments to be reported were designed to serve three principal purposes: (1) to provide further evidence for the existence of displacement-effects and their decay in time; (2) to provide further data on the distance-paradox, particularly the ascending portion of the curve which is not evident in the data of Jones and Bressler, and (3) to test a new deduction from the Köhler-Wallach model concerned with the effect of satiation on the variability of medial-plane judgments.

EXPERIMENTAL DESIGN

Experiment 1. The data of experiments on figural after-effects in vision relate to changes in the distance-relationships among simultaneously perceived contours (the simplest judgment possible being the comparison of the distance-relations among three parallel lines). Since the problem of interference makes this method inadvisable in audition, a method of single stimuli was employed. The *O*s were required to make the same judgment in all phases of the investigation, namely, to place a sound so that it appeared to lie in the medial plane. A correlate for mediality must therefore be derived from the Köhler-Wallach model. Köhler and Wallach make no statements about the perception of absolute locality, but, if a functional framework supplied by stimulation in other modalities or by anatomical boundaries is postulated, a correlate for mediality can be specified: *A stimulus will appear medial if its figure-currents interact equally with the framework to left and right.* In a homogeneous cortical medium this condition reduces to a symmetrical distribution of figure-currents about the center of the projection-area. If the impedance is greater on one side than on the other, there will be less interaction on that side and the stimulus will appear to be shifted to the side of the lower impedance.

We may ask how this model accounts for the variability of judgments of the medial plane. Fig. 1 will help in answering this question. Stimuli falling somewhere in the neighborhood of the medial plane but not exactly medial will give rise to asymmetrical interactions which are not noticeably different. We may assume that stimuli falling between the two points (*P*), one on each side of the medial plane (*M*), will fulfill this condition while stimuli falling outside this region will not. If it is asked what will be the effect of satiation at the medial plane, a new deduction from the Köhler-Wallach model follows. Satiation at *M* should result in a reduction of the spatial difference limen at that point. If we indicate a

region of increased impedance (shaded area) resulting from prolonged stimulation at M , we see that the figure currents of stimuli to the left of the medial plane encounter more impedance to the right, while those to the right of the medial plane encounter more to the left. The points P , which previously were indiscriminably different from M , now are noticeably different, and two new points (P') mark the boundaries of the functionally medial region. The first experiment was designed to test this deduction.

Procedure. The experiment was conducted in a sound-proof room. The O sat at a distance of 200 cm. from the center of a horizontal track perpendicular to his line of regard, with his head fixed in position by a bite-board. Mounted on the

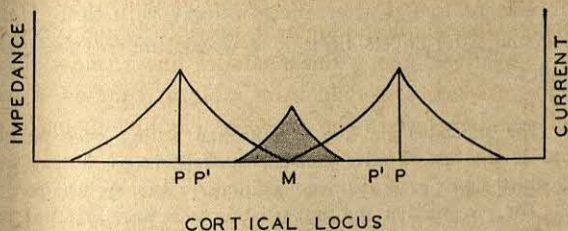


FIG. 1. HYPOTHETICAL DISTRIBUTION OF FIGURE-CURRENTS AND INCREASED IMPEDANCE

(M , medial plane; P , just noticeably non-medial locus before satiation; P' , just noticeably non-medial locus after satiation; the shaded area is a region of increased impedance produced by prolonged stimulation at M .)

track, which was 100 cm. long, was a Jensen VH-91 hypex horn. This speaker could be positioned on the track not only by E but also by O , who was provided with a remote-control knob connected to the speaker by a gear-and-shaft system. The apparatus and O were positioned symmetrically with respect to the walls of the room to reduce differential effects of sound-reflection from the walls. The stimulus, an unfiltered broad-band noise, constant in intensity and not much above threshold level, was generated by an apparatus located in an adjoining room. The same sound was used both for satiation and for determinations of the medial plane. The experiments were conducted in total darkness.

Ten graduate students in psychology served as O s. Each O was tested in two sessions, which were scheduled one week apart. In the experimental session O made 10 medial-plane settings (by the method of adjustment) both before and after a 2-min. period of satiation. The control session differed from the experimental only in that a 2-min. silent interval was substituted for the satiation period. Half the O s were tested first under the control condition, and the other half were tested first under the experimental condition. In order to minimize the satiation resulting from the testing series, the O s were encouraged to make the medial-plane determinations quickly and the sound source was turned off promptly following each setting.

Experiment II. The effect of satiation at points removed from the medial plane on subsequent determinations of the medial plane may be deduced from the same assumptions as the effect of satiation in the medial plane. A physically medial stimulus which encounters an asymmetrical distribution of impedance will appear shifted to the low-impedance side. If a region to the right of the medial plane is satiated, a physically medial tone should appear shifted to the left, and some point previously perceived to the right should now be accepted as medial. On the assumptions of Köhler and Wallach,¹³ the extent to which the medial plane is displaced should first increase as a function of the distance of the satiated region from the true medial plane, and then decrease (distance-paradox). By requiring *O* to make several settings after satiation, it is possible to study another property of figural after effects—decay with time. Hammer's results with visual after-effects indicate a decay-time of the order of one or two minutes.¹⁴

Procedure. The apparatus and the experimental setting were the same in this experiment as in the first. Twelve graduate students in psychology served as *O*s, with each *O* tested under each of three conditions in sessions separated by intervals of one week. The pre-satiation and post-satiation tests were identical with those of the first experiment. The satiating stimulus, which was again 2 min. in duration, appeared either 10, 20, or 40 cm. to the right of the mean of *O*'s pre-satiation settings of the medial plane. Two *O*s were tested in each of the six possible orders of the three conditions.

Experiment III. Experiment III was essentially a repetition of Experiment I. The procedure was the same, with 36 undergraduates serving as *O*s. Essentially the same apparatus was used for Experiment III, but several changes should be noted: An improved track was employed, with a string-control replacing the gear-and-rod drive, the effect being to eliminate the slight noise made during settings by the old apparatus; a low-pass filter was introduced into the circuit of the noise-generator to reduce the possibility of judgments made on the basis of intensity; the length of the track was increased to 150 cm.; and the experiment was conducted in the anechoic room of Mezes Hall.

Experiment IV. This experiment, a repetition of Experiment II, was conducted in the anechoic room of Mezes Hall with the improved apparatus of Experiment III. There were four loci of satiation—5, 10, 20, and 40 cm. to the right of the medial plane. The locus of the medial plane was defined, as before, by the mean of the first 10 settings. Of the 36 *O*s who served in Experiment III, 24 were chosen (on the basis of minimum variability) to fill the 24 possible orders of the four conditions of the fourth experiment.

¹³ Köhler and Wallach, *op. cit.*, 269-357.

¹⁴ E. R. Hammer, Temporal factors in figural after-effects, this JOURNAL, 62, 1949, 337-354.

RESULTS AND DISCUSSION

The results of Experiments I and III, shown in Table I, are presented in terms of the variability of initial and terminal medial-plane settings under the two conditions. The measure of variability was the standard deviation of the series of 10 determinations. All tests of significance were made by Wilcoxon's non-parametric method for paired replicates.¹⁵ In Experiment I the mean SD before satiation for the experimental session

TABLE I

RESULTS OF EXPERIMENTS I AND III

Standard deviations of medial-plane settings (in cm.). The significance of differences was tested by Wilcoxon's method for paired replicates.

Experiment	Condition	Mean SD		Diff.
		Pre-satiation	Post-satiation	
I	Experimental	6.84	4.53	2.31*
	Control	6.39	5.50	.89†
	Diff.			1.42†
III	Experimental	7.10	5.77	1.33*
	Control	7.71	7.27	0.44
	Diff.			0.89
* $P < .01$	† $P < .02$	‡ $P < .05$		

was 6.84 cm., and the mean SD following satiation was 4.53 cm., the difference being significant beyond the 1% level of confidence. In the control session, too, there was a significant decline in the mean SD (5% level) despite the fact that there was no interpolated stimulation. The decline in the experimental session was, however, larger than the decline in the control session, the difference being significant beyond the 2% level of confidence. The results of the third experiment show the same general pattern as those of the first—the reduction in variability was significant under the experimental but not under the control condition. The slight difference in the results of the two experiments probably is due to the fact that the Os of the first experiment were graduate students in psychology who were more practiced in psychophysical judgments than were the undergraduates of the third. (The initial variability scores were somewhat lower for the graduate group.) In any event, the results of the two

¹⁵ F. Wilcoxon, *Some Rapid Approximate Statistical Procedures*, American Cyanamid Co., 1949, 1-16.

experiments support the hypothesis that satiation at the medial point should reduce the difference limen at that point.

The results of Experiments II and IV were first analyzed in terms of the mean difference between pre- and post-satiation settings on each trial for each condition. In Experiment II, the mean shifts for the first trial of the 20-cm. condition, and the first, fourth, and fifth trials of the 40-cm. condition were significantly different from zero. In Experiment IV, the second trial of the 10-cm. condition and the first and second trials of the 40-cm. condition showed significant shifts. These results suggest

TABLE II

ANALYSIS OF VARIANCE OF MEAN DISPLACEMENTS ON TRIALS 1 AND 2 OF EXPERIMENT IV

Source	Sum of squares	df	Variance	F
Satiation-loci	631.37	3	210.62	7.45*
Os	1832.94	23	79.69	2.82*
Residual	1950.30	69	28.27	
Total	4415.11			

* Significant beyond the 0.1% level of confidence.

the validity of the dual hypotheses that greater shifts should be found for early trials and for satiation loci more removed from the medial plane. In order to examine the latter hypothesis more closely, an analysis of variance was performed on the mean shifts for the first two post-satiation

TABLE III

RESULTS OF EXPERIMENTS II AND IV

Mean displacements (in cm.) for the first two post-satiation trials

Experiment	Locus of satiation (in cm.)			
	5	10	20	40
II	—	—	—	—
Diff.	—	0.57	3.39	4.01
IV	—	—	2.82	0.62
Diff.	-1.33	2.46	1.94	5.90
		3.79*	-.52	3.96*

* Significant beyond the 2% level of confidence by the *t*-test; $\sigma_M = 1.53$ cm. (based on residual error variance).

trials for each *O* in each condition. The results of this analysis for Experiment IV are presented in Table II. In Experiment II the only significant *F*-ratio was that for *Os* (significant beyond the 5% level). In Experiment IV the *F*-ratios for both *Os* and satiation-loci were highly significant (beyond the 0.1% level of confidence. The difference between these results and those of Experiment II is probably a function of the

improved apparatus, the use of an anechoic chamber, and the greater number of Os.

The mean displacements for the first two trials of the post-satiation periods, together with the differences between the means, are presented in Table III, and the mean displacements for each condition of satiation are plotted in Fig. 2. The mean displacements on all 10 trials following satiation at 40 cm. in Experiment IV are plotted in Fig. 3. The curve shows a distinct decay with time; the mean displacements for the first

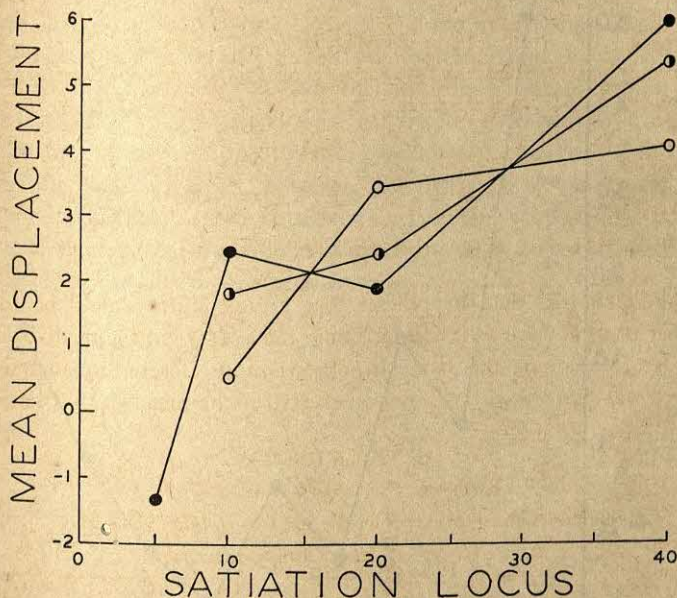


FIG. 2. DISPLACEMENT OF THE MEDIAL PLANE FOLLOWING SATIATION AT POINTS REMOVED FROM THE MEDIAL PLANE

The loci of satiation were 5, 10, 20, and 40 cm. to the right of the mean pre-satiation setting of the medial plane. Displacements are given in cm. The open circles are for Experiment II (12 Os), the solid circles for Experiment IV (24 Os), and the half-solid circles the two experiments combined.

two trials deviate significantly from zero (beyond the 1% level of confidence) and none of the other trials shows a significant after-effect.

Although there is some disagreement between the results of the second and fourth experiments, the general pattern is quite clear. Both experiments bear out the hypotheses that satiation at points removed from the medial plane should produce shifts in the settings of the medial plane toward the stimulated side, that the effects should increase as the dis-

tance of the satiation region from the medial plane increases, and that the effects should fade rapidly in time. The distance-paradox is, in fact, only partially demonstrated by these data (*i.e.*, only the rising portion of the

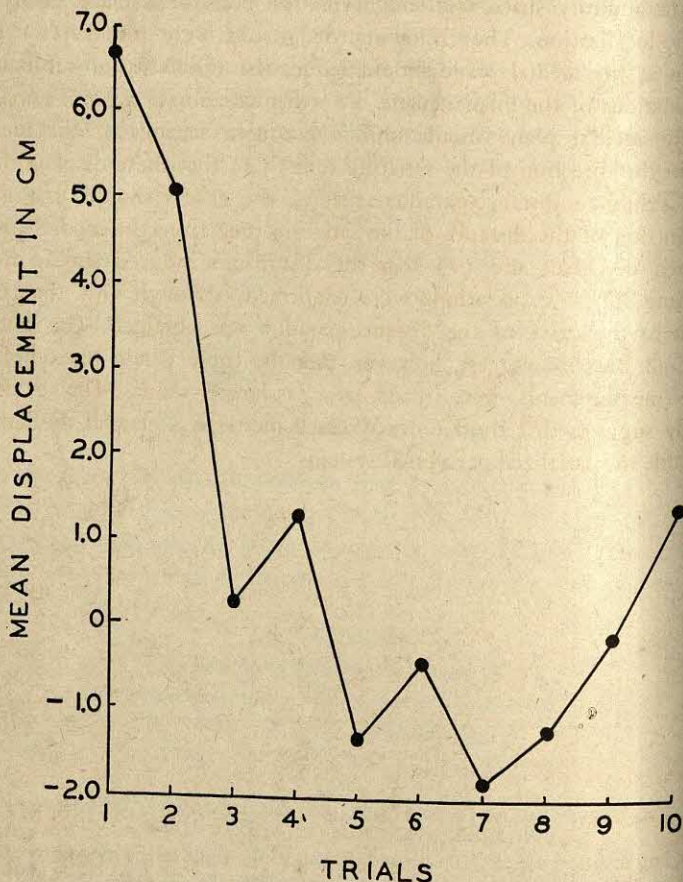


FIG. 3. DECAY OF AFTER-EFFECTS FOLLOWING SATIATION AT 40 CM. IN EXPERIMENT IV

curve appears). This is the crucial part of the curve, however, since Jones and Bresslar already have demonstrated the decline in the magnitude of the effects between 60° and 90° .

SUMMARY

If the Köhler-Wallach theory of figural after-effects is conceived as a general theory of neural organization applicable to all perceptual dimen-

sions which have a placewise representation, it follows that prolonged stimulation in any such dimension should yield alterations in spatial relationships. In the experiments here reported evidence of figural after-effects in auditory space was sought on the basis of a place theory of auditory localization. The following deductions were tested: (1) that satiation in the medial plane should reduce the variability of subsequent determinations of the medial plane; (2) that satiation at a point removed from the medial plane should shift subsequent settings of the medial plane in the direction of the satiating tone; (3) that these displacements should exhibit a distance-paradox, that is, the effects should first grow as a function of the distance of the satiating tone from the medial plane, and then decrease; and (4) that the distortions induced should decay with time. These expectations were confirmed, although only the rising portion of the curve of the distance-paradox was obtained. The data of Jones and Bresslar suggest, however, that the curve would be completed if the measurements were made over a larger range. The evidence strongly suggests that the Köhler-Wallach model is a general mechanism applicable to spatialized perceptual systems.

THE PRESTIGE AND IMPACT OF VARIOUS PSYCHOLOGISTS ON PSYCHOLOGY IN AMERICA

By SAMUEL W. FERNBERGER, University of Pennsylvania

During the winter of 1952-1953, the writer did a pilot study at the University of Pennsylvania in an effort to measure the prestige and impact of various psychologists upon present-day psychology. This attempt consisted in asking graduate students in psychology to check a list of names, indicating by one check the familiar names, and by two checks the names of those whose special fields of competency were known to them. The sample was small and localized in a single institution, but the results seemed of sufficient interest to warrant enlarging the study.

Letters were sent to chairmen in Departments in 119 institutions offering graduate work in psychology, asking them to cooperate in the study.¹ Favorable replies were received from 97 of them and the lists were sent in September, 1953. The returns were closed on January 15, 1954, inasmuch as it was desired that the first year students should be relatively unexposed to graduate instruction. By this date, returns had been received from 93 institutions located within the United States as follows: East, 34; South and Southwest, 24; Middle West, 21; Rocky Mountains, 7; West Coast, 6; and Hawaii, 1. A total of 2288 graduate students followed instructions and their returns are the basis of the present study.

Psychologists listed. That the selection of names should be objective, the following criteria were employed: (1) persons holding the presidency of the American Psychological Association from its founding in 1892 to 1925; and (2) persons contributing to one of the first three volumes of *History of Psychology in Autobiography* were listed.² From the latter, the *History*, we obtained the names of European psychologists and of some Americans who had not held the presidency of the Association prior to 1926. To the combined list thus obtained, four names were added: two from America, Titchener and Witmer (neither of whom was ever president of the Association but both of whom had made important contributions to psychology), and two from Europe, Ebbinghaus and Wundt (who were added

* Accepted for prior publication March 22, 1954. This study was aided by a grant from the Committee on the Advancement of Research of the University of Pennsylvania. The author's thanks are due Edwin G. Boring and Karl M. Dallenbach for their encouragement and many suggestions, and to the Departmental Chairmen and Graduate Students whose cooperation made this study possible.

¹ The institutions listed in *Amer. Psychol.*, 7, 1952, 64-83.

² *Op. cit.*, Carl Murchison, ed., Vol. 1, 1930; Vol. 2, 1932; Vol. 3, 1936.

to act as 'anchors' for comparison with other foreign psychologists). A list of 67 names was thus compiled. The oldest person on the list was Wundt, born in 1832, and the youngest was Piaget, born in 1896.

Procedure. The graduate students were asked to indicate (1) the institution that conferred their undergraduate degree, (2) the institution at which they were then doing graduate work, and (3) their present status, *i.e.* whether just beginning graduate work, or if not, the number of years of graduate study completed.

Instructions: The instructions were as follows: "The names of some psychologists are listed below. Read the list carefully and *single check* the names which are familiar to you, and *double check* each name if you can also state the specific field of the individual's contribution to Psychology."

Actually, in the writer's opinion, the mere familiarity with a name is of little consequence for the present study. For example, before the name of Walter D. Scott, one person wrote the word 'Sir.' It would seem that a minimal criterion of present impact should be, at least, a knowledge of the field of an individual's contributions.

It should be noted that there is the possibility that some of the names may be erroneously recognized. Though the list contained the first names and middle initials of all within it, some confusion may have arisen because the respondents were unfamiliar with the first names. As cases in point, the annual Directories of the American Psychological Association contain 13 Morgans, 12 Katzs, 9 Scotts, 4 Sterns, 4 Marshalls, 3 Hunters, 3 Sanfords, 2 Angells, 2 Baldwins, 2 Cattells, and 2 Thorndikes. It is difficult to see how this possible source of error could have been avoided and also how to estimate its extent.

Results. The criterion of a man's impact is knowledge of the field of his contribution rather than mere familiarity of his name. For purposes of completeness, however, Table I includes both categories. The rank correlation between rankings of the name recognized and rankings of field known is very low, having a rho of only 0.185. The names are arranged in the rank order of the double check, or knowledge of the field of contribution. The highest in rank order is E. L. Thorndike, whose field was known to over 90% of those responding. Watson, Terman, and Wundt follow closely behind and in that order. The relationships of known fields and rank order is in the form of an asymmetrical ogive.

Quartile ranking. The names seem to fall into four groups of almost equal size. The top 18 in the distribution had the field known by more than 75% of the respondents. The second group ranges from 75 to 25% and contains 15 names. The third group ranges from 25 to 10% and consists of 16 names. The last group, of less than 10% consists of 18 individuals. Although not exact, these groups will hereafter be designated as quartiles for purposes of convenience.

The upper quartile consists of 14 American psychologists—Thorndike, Watson, Terman, Gesell, Dewey, Hull, James, Thurstone, Titchener, Boring, Seashore, Wood-

worth, Tolman, and Cattell. In this group Wundt and Ebbinghaus are Germans, Spearman is British, and Janet is French.

Contrasted to this, the lowest quartile, of whom field is known, contains only the names of eight Americans—in order Royce, Bentley, Bryan, Elliott, Baird, Scott, Fullerton and Marshall—while it contains the names of ten Europeans—in order, Höfding, Michotte, Gemelli, Ferrari, Drever, Groos, Kiesow, Bourdon, Heymans, and Ziehen. The Europeans with ranking between 75 and 25% are McDougall (both European and American background), Piaget, Morgan, Katz, Stern, Münsterberg (again both European and American background), Stumpf, and Burt. Those in the third quartile of 25 to 10% are Zwaardemaker, Claparède, Piéron, and Thomson.

It is somewhat difficult to account for these differences in the relative standing of the Europeans. Some, like Wundt, Ebbinghaus, and Janet are historic figures—men who developed psychology as a science in its earliest stages. They have developed larger points of view or special fields. The same may be said of Thorndike, Watson, Terman, Gesell, Dewey, James, and Titchener in America, all of whom are in the upper quartile. Seashore and Cattell are probably better known for their experimental contributions among the older group. Hull, Thurstone, Boring, Woodworth, and Tolman of the younger members from America in the upper quartile have contributed largely to method and to point of view. The same may be said of Spearman.

On the whole, the European psychologists come out rather badly in the rank orders. Four (Wundt, Ebbinghaus, Spearman, and Janet) are in the upper quartile. McDougall and Münsterberg may well be considered Americans. This leaves six Europeans in the second quartile (Piaget, Morgan, Katz, Stern, Stumpf, and Burt). Only three foreign psychologists (Zwaardemaker, Claparède, and Piéron) are in the third rank order group (25-10%). In the lowest quartile, with field known less than 10%, 10 of the 18 psychologists are Europeans (Höfding, Michotte, Gemelli, Ferrari, Drever, Groos, Kiesow, Bourdon, Heymans, and Ziehen). Of the entire list of European psychologists, Spearman, Morgan, Burt, and Drever are British, and their contributions have been in English. The important work of Ebbinghaus, Piaget, and Katz has been translated into English. Most of the written contributions of the other foreign psychologists have been in the languages of their countries—French, German, Italian, Dutch, and Danish. It may be believed that one of the contributing factors to the rank orders, and especially the relatively poor showing of some of the foreign psychologists is the inability of American graduate students to read foreign languages. Another contributing factor may be the relative non-availability of foreign language books and scientific journals in the libraries of some American institutions.

Master-pupil relationship. A factor of undoubted influence upon the ranking is the number of important psychologists who have been under the instruction of some 'master.' The analysis, for some 102 pupils in relation to their 'masters' will be found in the article by Boring and Boring.³ All of the 'pupils' are psychologists who have been 'starred' in

³ M. D. Boring and E. G. Boring, Masters and pupils among American psychologists, this JOURNAL, 61, 1948, 526-534.

American Men of Science.⁴ Among the names in the present study, the frequency of 'master' relation is as follows: Hall, 13; Wundt, 11; Münsterberg, 9; Angell, 8; Woodworth, 7; Cattell, 6; Titchener, 6; Thorndike, 6; James, 4; Stumpf, 3; Seashore, Terman, Boring, Dodge, 2 each and Pillsbury, Watson, Dewey, Sanford, Dunlap, Judd, Witmer, and Tolman, 1 each. Hence, 12 of the names in the upper quartile are in the 'master' list; 5 in the second quartile, 5 in the third quartile, and none in the fourth quartile.

Productivity. There can be little doubt but that the quantity of production, which leads to publication, has an effect upon the impact on present-day psychology. In a recent article, Dennis emphasized this fact.⁵ Thorndike's bibliography runs to 507 titles. Wundt, Hall, James, and Titchener, in that order, have bibliographies of more than 240 titles. For most of the individuals under consideration in the present study, it was a simple matter to determine the size of their bibliographies by referring to the *Psychological Register*⁶ up to 1932 and to *Psychological Abstracts* subsequent to that date.

The results are not without interest. The average bibliography of the psychologists in the upper quartile is 187 titles. In the second quartile the average drops to 143 titles. The decline continues with an average of 127 titles for the third quartile, and an average of 83 for the last quartile. In the third group, Piéron has 469 titles. If he is eliminated, the average of his group is reduced to 101 titles. In the lowest group, the elimination of Gemelli with 394 titles lowers the group-average to 61 titles. It certainly seems that the number of times one's name appears in print has an influence on one's prestige and impact on the science.

Volume of publication alone does not, however, tell the entire story. Quality is an important factor. This is evident when one considers the case of Ebbinghaus, who, though rating fifth in rank order of 'known fields,' has a very small bibliography. To quote Dennis, "Ebbinghaus is decidedly the low man (of his group), his bibliography containing only 13 items."⁷ It appears, however, that in Ebbinghaus' case, his bibliography is a poor index of his actual activity as a scientist. . . . While high productivity is by no means a guarantee of eminence, it would seem that the majority of distinguished psychologists have been remarkable for the quantity as well as the quality of their scientific output."⁸ In this connection, it may be mentioned that Gregor Mendel's bibliography consists of only two titles and both were published in a small, local, and obscure journal.

⁴ S. S. Visser, *Scientists Starred in American Men of Science, 1903-1943, 1947*, 141-143. Cf. also this JOURNAL, 52, 1939, 279.

⁵ Wayne Dennis, Bibliographies of eminent psychologists, *Amer. Psychol.*, 9, 1954, 35-36.

⁶ *Op. cit.*, Carl Murchison, ed., Vol. 3, 1932.

⁷ This is an error as David Shakow (Bibliography of Hermann Ebbinghaus, this JOURNAL, 42, 1930, 516-518) lists 44 items.

⁸ Dennis, *op. cit.*, 35 f.

As was pointed out some years ago, relatively few psychologists publish in any quantity.⁹ The study referred to was made of the published contributions for the 5-yr. period, 1932-1936, of all the full members of the American Psychological Association—then numbering 587. Among these, 126 published nothing during this period, while almost one-half did not average one paper a year.

In addition to number of publications there are other factors pertinent to our problem that are very difficult if not impossible to evaluate: viz. quality, field, and method. Were the contributions confined to narrow fields or were they diversified? Were they accepted at the time of their appearance; have they withstood the test of time? Did they open up new points of view or new fields? Did they develop new techniques or new methods? Above all, are they of general interest to present-day graduate students in the United States? The lack of interest in certain fields of psychology in the United States may be the principal factor in the relatively poor showing of the European group. Though many of them have made important contributions to psychology, their contributions are not in fields that are of present interest here. The same, of course, is true of the Americans who stand low in the rank order. Some of these, presidents of the American Psychological Association during its early years, were not primarily psychologists. Marshall, an architect, and Fullerton, a philosopher, are examples.

Temporal effects. It is possible, from a comparison of the results of this and of Visser's study, to determine the effects of time upon the ranking of a number of psychologists whose names were used in this study.¹⁰ Visser gives the rank orders of frequency of votes in the initial 'starring' of psychologists appearing in the first edition of *American Men of Science* in 1903 in which the top 50 scientists in every field were elected by their colleagues. Of the 51 psychologists so designated, 27 appear in the present list. Table II gives the rank orders in 1903 and in this study, 51 years later. Obviously all of this group are Americans.

As this Table shows, only 6 of the 27 listed have gone up in relative rank. The greatest gain is by Woodworth. Thorndike and Seashore are tied for second place gain and they are followed in order of gain by Franz, Washburn, and Dewey. Dewey, 44 yr. old in 1903, lived a long and productive life, dying in his 93rd year in 1952. Thorndike and Franz were the youngest of the group, with Washburn only 3 and Woodworth only 5 yr. older. All of these psychologists continued to produce either up to the present time or until their recent deaths. Obviously, those at the bottom of the 1903 list had a handicap for gain, while those near the top of the early list had a handicap for loss of prestige. For those who have gained, Dewey and Thorndike were both near the center of the 1903 rank order and the others were near the bottom.

⁹ S. W. Fernberger, The scientific interests and scientific publications of the members of the American Psychological Association. Inc., *Psychol. Bull.*, 35, 1938, 261-281.

¹⁰ Visser, *op. cit.*, 143.

Of those who have lost prestige during these 51 yr., as indicated by rank order, Fullerton, Marshall, Royce, and Bryan show the greatest decline, in that order. Three of these were philosophers rather than psychologists, in the present-day sense. Bryan made one magnificent contribution to psychology but, in later years, went into academic administration. The next group to lose prestige were, in order, Baldwin, Calkins, and Sanford. Here the situation seems to be different from those just mentioned. All contributed to psychology but their points of view or experi-

TABLE II
RANK ORDERS IN 1903 AND 1954

Names	Rank Order		Gain or loss	Age 1903	Names	Rank Order		Gain or loss	Age 1903
	1903	1954				1903	1954		
James	1	9	-8	61	Stratton	15	38	-23	38
Cattell	2	17	-15	43	Thorndike	16	1	+15	29
Münsterberg	3	27	-24	40	Marshall	20	66	-43	51
Hall	4	22	-18	59	Judd	21	39	-18	40
Baldwin	5	44	-39	42	Angell	22	24	-2	36
Titchener	6	11	-5	36	Witmer	23	37	-14	36
Royce	7	50	-43	48	Warren	25	47	-22	36
Ladd	8	36	-28	61	Dodge	27	46	-19	32
Dewey	9	7	+2	44	Seashore	29	14	+15	37
Jastrow	10	28	-18	40	Pillsbury	37	45	-8	31
Sanford	11	41	-30	44	Washburn	42	34	+8	32
Calkins	12	42	-30	40	Woodworth	43	15	+28	34
Bryan	13	54	-41	43	Franz	44	35	+9	29
Fullerton	14	61	-47	44					

mental results have not remained important, apparently. One might have expected that Baldwin, for his early work in mental development, and Calkins for her system of self psychology, might have remained historical figures, but this does not seem to be the case. Sanford dropped psychology to go into academic administration.

In the middle group are Ladd, Münsterberg, Stratton, and Warren. The first three, apparently, are less well known today than formerly. One might expect Münsterberg to be an historical figure for those students interested in industrial applications, and Ladd as the writer of an early scientific textbook. Warren's greatest contribution to psychology was the development of psychological journals, and the names of journal editors are soon forgotten.

In the group with still smaller loss of apparent prestige are, in order, Dodge, Hall, Judd, Jastrow, Cattell, and Witmer. Hall has relatively well maintained his position in spite of being one of the older psychologists in the 1903 list. Both he and Witmer are more or less historical figures (even though Witmer is still living) with prestige because they both initiated work in psychological fields which are still of great importance in the present day science. The same may be said of Cattell and Judd, although both contributed important experimental results as well. Jastrow is probably remembered because of his popular and controversial publications.

Those who have best maintained their prestige over the 51-yr. period are Angell, James, Pillsbury, and Titchener, by showing smallest loss but no gain in the rank orders. Of these, the most remarkable is James, who headed the 1903 list and who was 61 yr. of age at that time. His contributions to both psychology and

philosophy seem to have withstood the passage of time. Titchener has also maintained his position, although he was sixth in the 1903 list. Both Angell and Pillsbury had the advantage of being low in the original list. The case of Angell is something of an enigma to the writer, because he also went into academic administration but relatively maintained his prestige position as a psychologist.

Sectional effects. There seemed to have been no influence exerted by the place of graduate work of the respondents. The analyses were made in lots of 300 in the order received and irrespective of the institution of origin.

TABLE III

RELATIVE PERCENTAGES OF KNOWN FIELD AT BEGINNING OF GRADUATE INSTRUCTION AND AFTER AT LEAST THREE YEARS OF GRADUATE INSTRUCTION

Name	Grad. instruct.		Gain	Name	Grad. instruct.		Gain	Name	Grad. instruct.		Gain
	Begin-ning	3+ yr.			Begin-ning	3+ yr.			Begin-ning	3+ yr.	
Angell	28	72	44	Groos	0.2	6	6	Sanford	10	20	10
Baird	4	9	5	Hall	44	87	43	Scott	2	10	8
Baldwin	3	29	26	Heymans	1	3	2	Seashore	64	92	28
Bentley	2	19	17	Höfding	5	14	9	Spearman	61	95	34
Bingham	9	50	41	Hull	66	99	33	Stern	13	66	53
Boring	63	92	29	Hunter	10	64	54	Stratton	5	49	44
Bourdon	0.2	2	2	James	72	97	25	Stumpf	13	59	46
Burt	9	61	52	Janet	59	94	35	Terman	81	98	17
Bryan	2	22	20	Jastrow	14	55	41	Thomson	3	22	19
Calkins	6	31	25	Judd	9	35	26	Thorndike	82	99	17
Cattell	63	94	31	Katz	16	61	45	Thurstone	65	97	32
Claparède	4	31	27	Kiesow	1	4	3	Titchener	67	97	30
Dewey	64	94	30	Ladd	9	43	34	Tolman	62	97	35
Dodge	2	31	29	Marshall	0.4	3	3	Warren	3	25	22
Drever	2	6	4	McDougall	54	92	38	Washburn	9	48	39
Dunlap	10	57	47	Michotte	3	14	11	Watson	76	98	22
Ebbinghaus	73	97	24	Morgan	30	82	52	Witmer	9	41	32
Elliott	3	11	8	Münsterberg	18	55	37	Woodworth	64	96	32
Ferrari	2	3	1	Piaget	44	89	45	Wundt	82	97	15
Franz	9	48	39	Piéron	3	25	22	Yerkes	41	97	56
Fullerton	0.4	6	6	Pillsbury	7	27	20	Ziehen	0.4	3	3
Gemelli	2	3	1	Royce	7	20	13	Zwaardemaker	6	31	25
Gesell	72	97	25								

It is quite astounding how little variation there was in the rank order from lot to lot.

Institutional effects. Also it was hoped that an analysis could be made regarding the influence of the institution of the respondents' undergraduate work. This proved impractical because of the great diversification found for each graduate institution. For example, a group of 70 at one middle western institution reported undergraduate instruction in 44 institutions. One eastern graduate group of 35 reported 23 undergraduate schools. One on the Pacific Coast, with 60 graduate students, reported 24 undergraduate schools, and 78 graduate students from a southwestern institution had their undergraduate work in 43 different institutions.

Years of graduate study. A study was made of the effect of at least three years of graduate instruction and training. A total of 475 students checked

that they were 'just beginning' their graduate work, while 327 checked that they were in their fourth year, while some few wrote in fifth or even sixth. The results in terms of percentages of double checks, indicating knowledge of field, will be found in Table III. The first and second columns give the *percentages* of double check for the first and fourth years respectively, and the third columns indicate the percentage gain over this period of graduate instruction.

Table IV gives the rank orders for the total sample of 2288 respondents as compared with the rank orders for the first and fourth years. The names appear in the rank order of the total sample and the results for the first

TABLE IV
RELATIVE RANK ORDERS: TOTAL SAMPLE, BEGINNING FIRST YEAR,
AND AFTER AT LEAST THREE YEARS

Names	Total	Begin- ning	Three plus Years	Names	Total	Begin- ning	Three plus Years	Names	Total	Begin- ning	Three plus Years
Thorndike	1	2	3	Angell	24	24	24	Dodge	46	54.3	41
Watson	2	4	4	Katz	25	26	27	Warren	47	52.5	46
Terman	3	3	3	Stern	26	28	25	Piéron	48	50.5	47
Wundt	4	1	5.5	Münsterberg	27	25	31.5	Thomson	49	50.5	48
Ebbinghaus	5	5	7.3	Jastrow	28	27	31.5	Royce	50	40	50.5
Gesell	6	7	11.5	Hunter	29	32	26	Höfding	51	44	53
Dewey	7	11.5	15.5	Stumpf	30	30	29	Bentley	52	59.5	53
Hull	8	9	1	Burt	31	33	28	Michotte	53	48	54
James	9	6	11.5	Dunlap	32	31	30	Bryan	54	54.3	49
Thurstone	10	10	5.5	Bingham	33	35	33	Elliott	55	49	55
Titchener	11	8	7.3	Washburn	34	36.3	35.5	Baird	56	46	57
Spearman	12	17	14	Franz	35	36.3	37.5	Scott	57	57.5	56
Boring	13	15	18.3	Ladd	36	39	37	Gemelli	58	54.3	66
Seashore	14	11.5	18.3	Witmer	37	34	38	Ferrari	59	57.5	65
Woodworth	15	13	13	Stratton	38	45	34	Drever	60	59.5	59.5
Tolman	16	16	7.3	Judd	39	36.3	39	Fullerton	61	63.3	58
Cattell	17	14	17	Zwaardemaker	40	43	40	Groos	62	66.5	59.5
Janet	18	18	15.5	Sanford	41	30	50.5	Kiesow	63	61	61
McDougall	19	19	18.3	Calkins	42	42	42.5	Bourdon	64	66.5	67
Yerkes	20	22	7.3	Claparède	43	47	42.5	Heymans	65	62	63.3
Piaget	21	30	31	Baldwin	44	52.5	44	Marshall	66	65.3	63.3
Hall	22	21	23	Pillsbury	45	41	45	Ziehen	67	63.3	63.3
Morgan	23	23	23								

and fourth years will be found respectively in the second and third columns.

It can be seen from Table III that the effect of graduate instruction increased the percentages in the case of every name. The top rank orders for the 'just beginning' group pair Thorndike and Wundt, and then in order, Watson, Ebbinghaus, Gesell, James, Titchener, Hull, Thurstone, Dewey, Seashore, Woodworth, Boring, Cattell, Tolman, Spearman, Janet, McDougall, Piaget, and Yerkes, the percentages ranging from 82 to 41. At the bottom of the list for beginning graduate work, with less than a 1% check, are Bourdon, Fullerton, Groos, Marshall, and Ziehen. From the respondents at this early stage, with 5% or less, 13 are Europeans while 11 are Americans. It is interesting to note that the highest 19 group in the total sample, although the rank order is somewhat different.

The group beginning graduate work double checked ('known field') an average

of 17.3 names, while the more advanced graduate students double checked an average of 34.7 names. Forty names was the highest number double checked for the beginning group, while for the fourth year group, one respondent double checked 63 names.

For the group who have had at least three years of graduate work, the situation is different. There were 20 names for whom the field was known to over 90% of this group. In rank order they are, Hull, Thorndike, Terman, Watson, Ebbinghaus, Gesell, James, Thurstone, Titchener, Tolman, Wundt and Yerkes, all of whom are known to 97% of fourth year students. Then follow in order the other eight: Woodworth, Spearman, Cattell, Dewey, Janet, Boring, McDougall, and Seashore, with percentages ranging between 97 and 90.

From this point the percentages of double check for fourth year graduate students drops off rapidly. In the following order there are 13 names with percentages ranging from 89 to 50: Piaget, Hall, Morgan, Angell, Stern, Hunter, Burt, Katz, Stumpf, Dunlap, Jastrow, Münsterberg, and Bingham.

At the lower end of the rank order for the fourth year graduate students, there are 14 individuals whose field is known to less than 10% of the respondents. In descending rank order they are Höfding, Elliott, Scott, Fullerton, Groos, Baird, Drever, Kiesow, Marshall, Ziehen, Bourdon, Heymans, Ferrari, and Gemelli. Of these 14, 9 are Europeans and 5 are from the United States.

The gain in percentages from the just beginning graduate student to the one who has had at least three years of graduate instruction is of interest. In this comparison, it must be remembered that a high percentage by the younger group gives the individual less chance to gain.

Of the 14 psychologists who showed the greatest percentage gain (from 56 to 41), the rank order is Yerkes, Hunter, Stern, Burt, Morgan, Dunlap, Stumpf, Katz, Piaget, Angell, Stratton, Hall, Bingham, and Jastrow. It is these whose impact and prestige is most enhanced by the three years of graduate instruction. At the end of undergraduate instruction, the fields of most of these individuals seem to be relatively unknown. The next group, whose percentages fall between 39 and 30 as the result of three years of graduate instruction, are, in rank order, Franz, Washburn, McDougall, Münsterberg, Janet, Tolman, Ladd, Spearman, Hull, Thurstone, Witmer, Woodworth, Cattell, Dewey, and Titchener.

The changes in rank order from the overall sample, the beginning group, and the advanced group are apparent in Table IV. An inspection of this table indicates that there is curiously little change in the three lists. The consistency is more apparent than the differences. Thorndike, in first place in the overall group, goes down to second place in the other two distributions. For the younger group, Wundt holds the highest ranking. For the advanced group, the first place goes to Hull. Of the first 13 in the total sample, all are in the top ratings of both of the other groups studied, with the exception of McDougall, who comes from the second quartile to tie, for the advanced students with Boring and Seashore. The outstanding case is Yerkes, who ties for seventh place with the advanced group, having been in 22nd ranking for the beginners and 20th for the entire sample.

The same consistency is true for the lowest ranks, those who, for the total sample had less than 10% double check. The consistency in rank order is certainly greater than the variation.

Conclusions. Neither the impact of an individual psychologist on present-day psychology, nor his prestige seems to be dependent upon the age of his contributions. Among the highest ranking psychologists, in this study, are those whose major contributions are recent, as well as those who are historical figures. For an overall sample of American graduate students at all levels of graduate instruction, only relatively few individuals seem to have made a really lasting impression. The field of contribution of only 18, of our total list of 67 names, was known to 75% of the respondents, and only six more individuals were so reported in the percentage range from 75 to 50. These 24 individuals represent both Europeans and Americans, both old and present day psychologists. This seems a small portion of those who, at one time or another, had enough prestige to be elected president of the American Psychological Association, or to be invited to submit their autobiographies in a published collection. The active European psychologists and many of recent date seem relatively unknown to American graduate students. Part of this may be due to the fact that their publications are not in English.

The results show that three years of graduate instruction changes the prestige of some psychologists in a few cases, but the consistencies of rank order of those reported as just beginning graduate work and those who have had at least three years of such instruction are far greater than the differences in relative rank order.

There seem to be several factors that are the basis for this impact and prestige. One of these is certainly the amount of publication. In the groups from highest to lowest rank order quartiles, there is a consistent decrease in the average size of bibliography. A second factor is the quality of the published contribution. Ebbinghaus, who ranks from fifth to seventh in all three distributions, had a small bibliography. The third factor seems to be the kind of contribution. This can be, like Ebbinghaus, the opening of a new field; like Thorndike, the contributions to a number of fields; like Watson, the development of a new point of view; or like Spearman, the development of a new technique. Not only must there be these kinds of contribution, but also the present-day graduate students in psychology must still be interested in this contribution, either for its present usefulness or as an historical event. Watson, who ranks second in the total sample, presented a point of view that is still widely accepted; while Calkins, rated forty-two, also presented a basic point of view which has not maintained popular interest or acceptance.

AN ANALYSIS OF ERRORS IN RETROACTIVE INHIBITION OF ROTE VERBAL LEARNING

By JAMES DEESE and GEORGE W. HARDMAN, JR.
The Johns Hopkins University

The literature on retroactive inhibition is very large, yet there has been surprisingly little attention to the kind of errors responsible for the observed retroactive effects. This, to some extent at least, is a result of the emphasis upon numerical indices of the total amount of retroaction as the dependent variable in transfer experiments. Most theories of retroaction or transfer are designed primarily to predict the amount of retroaction or transfer rather than the nature of S's behavior which is responsible for the effect.

One frequently cited experiment on the analysis of errors in retroactive inhibition is that of McKinney and McGeoch.¹ These authors used a method of complete presentation with retention tested by written recall. Both immediate- and delayed-recall tests were given, though complete data are presented only for the delayed-recall tests. Unfortunately, the experiment involved group testing, and, as a result, the data are probably not as reliable as they should be. The result emphasized by the authors, and one that has found its way into almost every secondary source on the subject,² is that about one-quarter of the obtained inhibition in the experimental conditions could be accounted for by intrusions of interpolated material during the delayed-recall tests of original material.

Another set of studies on the analysis of responses during retroactive inhibition is that by Osgood,³ who presents a more or less complete analysis of errors in paired-associates experiments on retroaction. The results of his experiments will be discussed later; however, it should be pointed out here that he employed paced learning, the typical method for experiments in rote learning. In the present laboratory it has been discovered that much greater success can be achieved in the analysis of errors during verbal learning if the traditional pacing method is replaced by a method that allows the S unlimited response-time.⁴ That is to say, *different* results are achieved if unlimited response time is allowed. The present authors feel this

* Accepted for publication March 18, 1953.

¹ F. McKinney and J. A. McGeoch, The character and extent of transfer in retroactive inhibition: Disparate serial lists, this JOURNAL, 47, 1935, 409-423.

² For example, J. Deese, *The Psychology of Learning*, 1952, 203.

³ C. E. Osgood, Meaningful similarity and interference in learning, *J. Exper. Psychol.*, 36, 1946, 277-301; An investigation into the causes of retroactive interference, *J. Exper. Psychol.*, 38, 1948, 132-154.

⁴ Deese and F. H. Kresse, An analysis of errors in rote serial learning, *J. Exper. Psychol.*, 43, 1952, 199-202.

method to be generally superior in the analysis of errors since it more nearly exhausts the potential for responding in *S*.

The present experiments are concerned with an extension of the method of unlimited response-time to a further analysis of errors in retroactive inhibition. Data are reported for (1) anticipative learning by the method of serial exposure, (2) paired-associates learning in serial exposure, and (3) learning of connected discourse by serial exposure.

METHOD AND PROCEDURE

The *Ss* used were 98 undergraduate students at The Johns Hopkins University. No *S* had previously taken part in experimentation on verbal learning, and all were naïve about the nature of retroaction and the purpose of the experiment. All *Ss* were tested individually. Only one experimental session was used; relearning occurred immediately after interpolated activity. The material in all experiments was presented serially, though method of exposure differed from experiment to experiment. A total of twelve trials was given the experimental *Ss* and a total of eight trials to control *Ss*. The details of the separate experiments are given below.

Anticipative learning. Two groups of 10 *Ss* each were used in this experiment. Both groups practiced for four trials the original material which consisted of 12 two-syllable adjectives drawn from Melton's list.⁵ The experimental group followed this learning with practice on an interpolated list made up of 12 highly similar adjectives (the median of Melton's similarity rating to original material was 2.36) for four trials. Finally, this group practiced the original list again for four trials. The control group practiced an irrelevant task (counting *As* in a mimeographed prose passage) for a period of time equivalent to interpolated learning. The lists were exposed on a memory-drum. On the first trial of each list the syllables were exposed for 2 sec. with a 2-sec. blank space following each syllable. On all subsequent trials (including the first trial of relearning) the exposure apparatus was halted at the blank space before each syllable in order to give *S* ample opportunity to anticipate the next syllable. If *S* was unable to make a response, he so informed *E* and the next syllable was then presented.

Ss were given instructions appropriate for serial learning and tested on several preliminary items to make sure that they understood the procedure. In addition, they were informed that *E* was interested in the mistakes people make during learning; consequently, they were not to hesitate to respond even if they were not sure they were right. *Ss* were not forced to guess, however, and failures of response could therefore, be recorded.

Paired-associates learning. Three groups of 18 *Ss* each were used in this experiment. All *Ss* learned a series of sentences presented in paired-associates form. There were 12 sentences in original learning and 12 in interpolated learning. The first part of each sentence was presented as a stimulus-item, following which the concluding word of the sentence was presented as a response item. For example, one stimulus-item was *Arsenic is . . .*, and the response-item during original learning was *deadly*.

⁵ E. R. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 548.

The pairs were presented in a random order different for each trial (the randomization was based upon a latin square for the total number of trials in the experiment). The stimulus- and response-items were printed on 3 × 5 in. cards. On the first trial *E* exposed the items at about the rate of 2 sec. per card. On subsequent trials the stimulus-item was held before *S* and the response-item not exposed until *S* either responded or said that he could not respond. The response-item was then exposed and 2 sec. later *E* exposed the next stimulus-item. During interpolated learning one experimental group learned to associate the same stimulus-items used in original learning with a word very similar to the original response-item. For example, the response-item appropriate to *Arsenic is* in this interpolated condition was *poisonous*. For the same stimulus-items another experimental group learned to associate unrelated words. In the case of the stimulus-item given above, the appropriate response-item for this interpolated condition was *tasteless*. In all cases the response-term made an intelligible sentence when paired with the appropriate stimulus-item. A control group practiced counting *As* in a prose passage during the interpolated period. All groups practiced the original material for four trials during relearning.

Connected-discourse group. In this experiment two groups of 12 *Ss* each (one group was reduced to 11 because it was necessary to discard an illegible record) learned, by rote, passages consisting of connected discourse. Both groups learned an original passage. One of the groups learned a second passage during the interpolated period while the other practiced an unrelated sorting task. (The counting of *As* was not used as a control in this experiment because the duration of the interpolated period was very long, and the counting task became very annoying to *Ss* when carried out for protracted periods.) The sorting task consisted of cards with one cardinal compass point indicated by an arrow and the designation *North* placed at random. The cards were to be sorted into four piles. The task kept *Ss* occupied and interested, while at the same time it was not fatiguing or annoying. Both groups relearned the original passage. There were four trials on original learning, four on interpolated learning, and four on relearning. The serial method of learning was forced by the method of presentation. On the first trial of each passage *E* read the passage to *S* while *S* copied it. *S* then took a fresh sheet and attempted to write the passage word by word. If *S* made an error or came to a word he could not remember, *E* prompted him by giving him the correct word. *S* then wrote the correct word above his guess (if it were incorrect) or above a blank space (if he had guessed nothing). *E* never allowed *S* to go on without correction once *S* made a mistake. The two passages are given below. The first is the original passage and the second is the interpolated passage. Scored words (see explanation below) are in italics.

"The present view of the nature of energy is dictated by the atomic concept of matter according to which atoms have an organization consisting of nuclei intimately associated with positive charges of electricity called protons surrounded by negative particles of force designated as electrons."

"The contemporary outlook of the substance of organisms is determined by the cell theory of structure according to which protoplasm has a form made of a nucleus directly composed of small strands of material named chromosomes enclosed by clear masses of liquid termed cytoplasm."

Method of analysis. All responses were recorded verbatim. The errors were

analyzed into the following classes: (1) within-list errors, (2) between-list errors, (3) errors not directly from either list, and (4) failures of response. A within-list error was an item incorrect but from the same list. Between-list errors were responses appropriate to the other list. In addition, in the paired-associates experiment, between-list errors were classified on the basis of whether or not they were given to a stimulus-item with which it had been associated in the other list.

The connected-discourse experiment presented a more difficult problem, since in order to make such passages it was almost impossible to avoid words in common between the two lists. These words are largely prepositions, articles, and connectives, which—being impossible to classify—made it necessary to add a category of unclassifiable errors. Furthermore, the scoring was done only on the items (italicized above) which were different between the two passages.

Statistical treatment. Data for all parts of each experiment are presented below, though statistical analysis of the significance of mean differences is presented only for the relearning data. Since, however, the analysis of the data from relearning demanded a correction in terms of regression on scores during original learning, analyses of variance were performed on the data from original learning. In no case was there a significant difference during original learning. For each mean difference discussed below, the appropriate value of t (obtained from analysis of covariance) is given. The value of t is given rather than the value of F because each analysis concerns only two groups and t is much more familiar in the two-group case than is F .

RESULTS

Anticipative learning. Table I shows the mean errors in each class for original learning, interpolated learning, and relearning, in the anticipative

TABLE I
MEAN NUMBER AND PERCENTAGE OF ALL TYPES OF ERROR IN ANTICIPATIVE LEARNING

Condition	Within-list		Extra-list		Between-list		Failure of response		Combined Mean
	Mean	%	Mean	%	Mean	%	Mean	%	
Control O.L.	10.1	66.0	0.6	3.9	0.0	0.0	4.6	30.1	15.3
Exp. O.L.	11.0	58.5	2.4	12.8	0.0	0.0	5.4	28.7	18.8
Exp. I.L.	4.6	45.5	1.5	14.9	0.3	3.0	3.7	36.7	10.1
Control R.L.	2.0	83.3	0.1	4.2	0.0	0.0	0.3	12.5	2.4
Exp. R.L.	4.6	50.5	0.5	5.5	0.5	5.5	3.5	38.5	9.1

learning experiment. In each case the data from all trials under each condition have been combined. In addition, this table converts each class of errors into percentage of the total. From a comparison of the total errors during relearning in the experimental and control groups, it is easy to see that there was a considerable retroactive effect of interpolated learning ($t = 5.10$). The most striking fact, however, is that the errors

attributable to between-list confusion are extremely small even in the experimental group. For this group they account for only 5.5% of the total errors. Another striking fact is that the within-list errors are larger in the experimental condition than in the control condition ($t = 3.53$). Thus, in part at least, the effect of interpolated learning is to increase the frequency of confusion within the original material. Further, the number of failures of response is larger in the experimental group than in the control group ($t = 2.91$).

Paired-associates learning. The results of the paired-associates experiment are shown in Table II. Again, it is apparent that there is a retroactive

TABLE II
MEAN NUMBER AND PERCENTAGE OF ALL TYPES OF ERROR IN PAIRED-ASSOCIATES LEARNING

Condition	Within-list		Extra-list		Between-list		Failure of response		Combined Mean
	Mean	%	Mean	%	Mean	%	Mean	%	
Control O.L.	1.78	31.1	1.39	24.3	0.00	0.0	2.56	44.7	5.72
Exp. Syn. O.L.	1.78	29.1	1.72	28.2	0.17	2.7	2.44	40.0	6.11
Exp. Unrel. O.L.	1.33	24.7	1.67	30.9	0.11	2.1	2.28	42.3	5.39
Exp. Syn. I.L.	0.44	10.7	0.89	21.3	1.89	45.3	0.94	22.7	4.17
Exp. Unrel. I.L.	1.94	28.5	1.00	14.6	0.94	13.8	2.94	43.1	6.83
Control R.L.	1.50	62.8	0.61	25.6	0.00	0.0	0.28	11.6	2.39
Exp. Syn. R.L.	0.44	9.4	0.67	14.1	3.22	68.2	0.39	8.4	4.72
Exp. Unrel. R.L.	1.06	27.9	0.78	20.6	0.89	23.5	1.06	27.9	3.78

effect measured in terms of total errors. In the case of the comparison of the control condition with the interpolation of synonyms, the t is 5.85; in the case of control and interpolation of unrelated words the t is 2.65. There is, however, a striking difference between this experiment and the previous one. In the case of paired-associates learning, a very large portion of the observed retroactive effect is due to intrusions of items during interpolated learning. In the group that learned synonyms of the original response-items for interpolated material, the between-list errors account for 68% of the total errors. In the group that learned interpolated items unrelated to original response items, the number of between-list errors is smaller, though such errors still account for 24% of the total errors during relearning. Thus, between-list errors much more readily account for a significant portion of the retroaction effect in paired-associates learning than in anticipative learning. It turns out also, that, in paired-associates learning, the effect of interpolation seems to be to reduce the confusion

within the original list. In both the control-synonym comparison ($t = 2.58$) and the control-unrelated comparison ($t = 3.01$), the within-list errors are smaller in the experimental conditions than in the control conditions.

It should also be pointed out that the greater frequency of between-list errors found in the experimental conditions of paired-associates relearning is largely due to Ss giving responses appropriate in the other list. In the synonym group 96% of the between-list errors during relearning were appropriate, while in the unrelated group 77% of the between-list errors during relearning were appropriate.

Connected-discourse learning. The mean number of errors (and percents) in each category for the connected-discourse experiment are presented in Table III. It is immediately apparent that there is only a very

TABLE III
MEAN NUMBER AND PERCENTAGE OF ALL TYPES OF ERROR IN CONNECTED-DISCOURSE LEARNING

Condition	Within-list		Extra-list		Between-list		Failure of response		Unclassified		Combined Mean
	Mean	%	Mean	%	Mean	%	Mean	%	Mean	%	
Control O.L.	14.5	36	4.8	12	1.3	3	11.3	28	8.3	21	40.2
Exp. O.L.	13.5	41	1.7	5	0.5	1	8.7	27	8.4	25	32.8
Exp. I.L.	9.6	26	2.8	8	7.3	19	10.6	28	7.2	19	37.5
Control R.L.	8.7	39	2.3	10	1.0	5	6.2	28	4.0	18	22.2
Exp. R.L.	4.8	21	1.9	8	6.0	27	5.5	24	4.4	19	22.6

small difference between the control and experimental groups in total errors during relearning. This difference, which yields a t of 1.36, is clearly not significant. Despite the fact that there is no apparent retroactive effect on total errors during relearning, there appear to be some interesting differences in the patterns of errors between the control and experimental groups. It is apparent both in absolute frequency and percentage-wise that the experimental condition produced more between-list errors than did the control. The difference in mean number yields a t of 2.98. Perhaps it is remarkable that the control condition produced *any* between-list errors, since the control Ss never saw the interpolated list. The fact that between-list errors do occur in the control condition, however, is due to the high strength of certain interpolated words within the context of the original material. For example, the phrase *atomic concept* in the original material was sometimes replaced by *atomic theory* by control Ss; *theory* was a word appropriate to the interpolated material. Despite these 'chance'

occurrences of interpolated words, the experimental group clearly showed more between-list errors during relearning than did the control. This difference, however, is almost completely counterbalanced by a greater frequency of within-list errors in the control condition during relearning ($t = 2.64$). Thus, while it appears that between-list errors can be produced by the interpolation of connected-discourse learning like that used in the present experiment, it also appears that within-list errors can be reduced sufficiently to result in no net retroaction.

DISCUSSION

The major finding in the present study is concerned with the differences in patterns of error during relearning (and hence in retroaction) produced in the three experiments performed. In the experiment using anticipative learning, most of the increase in errors due to interpolation was the result either of within-list errors or failures of response. A very small portion of such errors was actually due to intrusions of interpolated responses during relearning; likewise, there was a very small residue of errors not assigned to any of these categories (extra-list errors). On the other hand, the between-list errors contributed a sizeable amount to errors during relearning in the paired-associates experiments. Likewise, in these experiments, interpolation appeared to reduce within-list errors. Finally, in the connected-discourse experiment, it was shown that, despite the lack of a *net* retroactive effect, there was a pattern of effects produced by interpolation like those in the paired-associates experiment. Between-list errors were increased by interpolation while within-list errors were reduced.

The importance of these findings is two-fold. First of all, they reinforce the conviction that retroactive inhibition is a complex phenomenon, the causes of which are several and variable with experimental conditions. The fact, for example, that the pattern of errors in relearning for anticipation is markedly different from that for paired-associates suggests that these two conditions of learning produce different kinds of retroactive effect. Secondly, the results of this experiment add weight to the notion that interference from interpolated learning at the time of relearning is not necessarily the most important cause of retroaction.

The idea that competition from interpolated responses at recall produces retroaction seems to apply fairly well to the case of paired-associates learning. As a more general notion, however, the present experiments suggest that the source of competition in retroaction is determined by the context of learning. For example, anticipative learning forces the learner to associate successive items with one another, so that a given item is always

within the context of the list in which it appears. In paired-associates learning, however, the task for the learner is no successive association. The stimulus-item itself, not the list in which it appears, is the primary determiner of context. By the nature of retroaction experiments, there are associated with any given stimulus-item two responses of approximately equal strength (though one is more recent than the other). Consequently, the context forces errors based on this common connection. The passages for the experiment on connected discourse were constructed in an attempt to maximize the possibility of confusion between the contexts of the two passages used. That the attempt was at least partially successful is indicated by the fact that the between-list errors were appreciably increased by interpolated learning.

In view of the context theory schematically stated above, a word needs to be said about the relation of the present experiment to the earlier ones cited in the introduction. McKinney and McGeoch reported that about one-quarter of the total errors produced in a delayed measure of retention (recall) could be attributed to interpolated responses.⁶ This was only true, however, for delayed recall (15 min. and 11 days); the exact figures for immediate recall were not presented. By inference from the published tables, however, it is apparent that between-list errors for immediate recall were negligibly small, as in the present anticipative experiment. The delay in recall must, therefore, have served to reduce the boundaries between the contexts of the two lists. This interpretation deserves further exploration.

In the paired-associates experiments a considerably higher proportion of between-list errors was found than appeared in Osgood's experiments using the same technique.⁷ This outcome is also predictable from the context hypothesis. The stimuli in Osgood's experiment were not 'meaningful'; that is to say, they had no connection with the response-item in S's histories. The stimulus-items were letter-pairs such as *tg* and the response-items were adjectives. In the present experiment the stimulus- and response-items, even in the case of appropriate between-list errors, made meaningful sentences (for example, *Arsenic is deadly* and *Arsenic is poisonous*), but the occurrence of a within-list error would not make sense (for example, a typical within-list error would result in the sentence, *Arsenic is empty*). Such was not the case in Osgood's experiment. Within-list errors were just as meaningful (or meaningless) as between-list errors. Thus, the errors were not nearly so liable to be determined by the context of the

⁶ McKinney and McGeoch, *op. cit.*, 409-423.

⁷ Osgood, *op. cit.*, 277-301, 132-154.

stimulus-items themselves, but were more typically associated with the set for the particular list being practiced.

SUMMARY

An analysis of the source of errors producing retroaction was performed on the data from three experiments: (1) learning by the method of anticipation, (2) learning by the method of paired-associates, and (3) learning of connected discourse by serial presentation. The results show that the source of retroaction depends upon the kind of learning and material used. In anticipative learning of lists of adjectives, practically all of the increase in errors in relearning due to interpolation was the result of either failure to respond or confusion of items within original learning. Only 5.5% of the total errors in relearning after interpolation could be directly attributed to the intrusion of interpolated material. In the case of paired-associates learning, a much larger proportion of the total errors in relearning could be directly attributed to intrusions of interpolated material. Depending upon the material used, the percentage varied from 24 to 68. In the case of paired-associates learning, the effect of interpolation was to reduce the within-list errors during relearning. The learning of connected discourse produced no retroaction at all measured in total errors. The net effect of interpolation in this case was to increase the number of intrusions of interpolated material during relearning and to decrease the number of within-list errors. The results of the present experiment and earlier ones were discussed in terms of the hypothesis that the source of associative errors in relearning depends upon context.

SUCCESSIVE AND SIMULTANEOUS DISCRIMINATION AS A FUNCTION OF STIMULUS-SIMILARITY

By EUGENE F. MACCASLIN, Human Resources Research Office

The relative difficulty of simultaneous and successive problems in two-choice discriminative situations has been studied in several recent experiments.¹ The problems differ in two important respects, and it is these differences to which we may look for an explanation of their relative difficulty. First, while the simultaneous problem can be solved in terms of the differential reinforcement of afferent components, the successive problem cannot. Thus, Spence's early theory of discriminative learning,² which dealt only with the consequences of differential reinforcement along specific afferent dimensions, could not handle the solution of successive problems, and extensive revision became necessary.³ Spence now treats the successive solution in terms of the concept of stimulus-compounding, and he is thus led to the conclusion that the successive problem must inevitably be more difficult than the simultaneous. Although some of the evidence which appears to contradict this conclusion⁴ may be circumvented by Spence in terms of still a third category of solution which is called "transverse patterning,"⁵ other evidence points to the reality of a successive solution ("configurational discrimination") based on a process more primitive than is implied by the concepts of compounding and transverse patterning.⁶

* Accepted for publication April 24, 1953. This paper is adapted from a doctoral dissertation submitted to the University of Texas. The author is indebted to Professor M. E. Bitterman, who directed the research.

¹ The series of investigations was initiated and some of the fundamental issues stated by Philip Weise and M. E. Bitterman, Response-selection in discriminative learning, *Psychol. Rev.*, 58, 1951, 185-195.

² K. W. Spence, The nature of discrimination learning in animals, *Psychol. Rev.*, 43, 1936, 427-449.

³ Spence, The nature of the response in discrimination learning, *Psychol. Rev.*, 59, 1952, 89-93; see also the comment by Bitterman, Spence on the problem of patterning, *Psychol. Rev.*, 60, 1953, 123-126.

⁴ Bitterman and Jerome Wodinsky, Simultaneous and successive discrimination, *Psychol. Rev.*, 60, 1953, 371-376.

⁵ Spence, The nature of the response in discrimination learning, *op. cit.*, 89-93.
⁶ D. C. Teas and Bitterman, Perceptual organization in the rat, *Psychol. Rev.*, 59, 1952, 130-140; Bitterman and J. V. McConnell, The rôle of set in successive discrimination, this JOURNAL, 67, 1954, 129-132; Wodinsky, M. A. Varley, and Bitterman, Situational determinants of the relative difficulty of simultaneous and successive discrimination, *J. Comp. & Physiol. Psychol.* (in press).

A second important difference between simultaneous and successive presentations, which is independent of the first, was stressed by Weise and Bitterman.⁷ Consider, for example, a black-white (*B-W*) discrimination, the stimulus-pairs of the simultaneous problem being *BW* and *WB* and those of the successive problem being *BB* and *WW*. Irrespective of how the solution of the successive problem is achieved, the solution of both problems depend upon the animal's ability to differentiate *B* and *W*. It is obvious, however, that opportunity for direct comparison of these stimuli is afforded by the simultaneous problem but not by the successive. Accepting the results of Saldanha and Bitterman as demonstrating the extreme difficulty (verging upon impossibility) of discrimination between very similar stimuli in the absence of opportunity for direct comparison,⁸ Weise and Bitterman predicted an increase in the relative difficulty of the successive problem with an increase in stimulus-similarity.

In general, then, differences in the difficulty of simultaneous and successive problems may be traced to at least two independent factors inherent in the structure of the problems—a component-noncomponent factor and a comparison-noncomparison factor—and the implications of each cannot be understood except in relation to the other. The experiment here reported was designed to demonstrate the rôle of opportunity for comparison of the stimuli to be discriminated. The hypothesis tested was that an increase in the similarity of the stimuli should increase the difficulty of the successive problem relative to that of the simultaneous.

Subjects. The *Ss* of the experiment were 52 experimentally naïve, male, Albino rats, bred in the University of Texas Laboratory from Wistar stock. The animals were 3-4 mo. old at the outset of the experiment.

Apparatus. A Lashley jumping apparatus was employed which had two 5 × 5 in. windows separated by a partition 1.5 in. wide extending 2 in. in front of them. There was a felt net 2 ft. below the windows into which an animal fell following an incorrect choice during training, and behind the windows was a feeding-platform containing a small food-cup at the center-rear to which the animal was admitted following a correct choice. The circular jumping-platform was equipped with a wire grid through which mild shock could be administered to break resistance to jumping. The entire apparatus was painted flat black. Seven different kinds of stimulus-cards were used in the experiment: (1) a small black circle (1 in. in diameter) centered on a white ground; (2) a large black circle (4 in. in diameter) centered on a white ground; (3) wide horizontal black-and-white stripes (the width of each stripe being 0.5 in.); (4) wide vertical black-and-white stripes (the width of each stripe being 0.5 in.); (5) narrow vertical black-and-white stripes (the

⁷ Weise and Bitterman, *op. cit.*, 188.

⁸ E. L. Saldanha and Bitterman, Relational learning in the rat, this JOURNAL, 64, 1951, 37-53.

width of each stripe being 0.25 in.); (6) homogeneous gray with a reflectance of 10%; and (7) homogeneous gray with a reflectance of 40%. Several duplicates of each type of card were necessary for the work.

Discriminative problems. Four discriminative problems were studied in the experiment. Two were simultaneous problems and two were successive. One of the simultaneous problems (I) and one of the successive problems (II) required a discrimination between relatively dissimilar stimuli; while the other simultaneous (III) and successive (IV) problems involved quite similar stimuli. Each of the problems was really a double problem—that is, the animal was required to learn discriminations between two pairs of stimuli concurrently. In Problem I the Ss learned simultaneous discriminations between the two circles differing in size and between the wide horizontal and vertical stripes. Both discriminations were known from previous work to be relatively easy—that is, the size difference between the circles and the directional difference between the stripes were large and readily

TABLE I
ILLUSTRATION OF THE FOUR PROBLEMS

Card-pair*	Problem				Rewarded response
	I	II	III	IV	
1	HV	HH	Vv	VV	right
2	VH	VV	vV	vv	left
3	LS	LL	Gg	GG	right
4	SL	SS	gG	gg	left

* H, wide horizontal stripes; V, wide vertical stripes; L, large circle; S, small circle; v, narrow vertical stripes; G, light gray; g, dark gray. HV means wide horizontal stripes in the left window and wide vertical stripes in the right window.

discriminable. In Problem II the Ss learned successive discriminations with the same stimuli as were used in Problem I. In Problem III the Ss learned simultaneous discriminations between light and dark gray and between wide and narrow vertical stripes. These brightness and size differences were small compared with the directional and size differences involved in Problems I and II. In Problem IV the Ss learned successive discriminations with the same stimuli that were used in Problem III. The four problems are illustrated in Table I. It should be noted that there were four possible variations (sub-problems) of each main problem. For example, in Problem I the Ss could be rewarded for responding to large circle and vertical stripes, large circle and horizontal stripes, small circle and vertical stripes, or small circle and horizontal stripes. Similarly, in Problem II the Ss could be reinforced for jumping left to horizontal stripes and large circles and right to vertical stripes and small circles, left to vertical stripes and large circles and right to horizontal stripes and small circles, and so forth. All four variations of each problem were used, the Ss in each group being randomly assigned to one or the other of the variations.

Preliminary training. The Ss were placed on a 24-hr. feeding schedule for a period of two weeks during which they were handled extensively, and they were then trained to respond to open windows in the jumping apparatus. The jumping-platform was moved as near as possible to the windows. On each of two days, prior

to daily feeding, each *S* was placed on the jumping-platform and allowed to move freely about from this platform through the open windows onto the feeding-platform where food was available. During the next four days this procedure was continued, with the distance of the jumping-platform from the open windows being gradually increased by 0.5-in. steps to the maximum of 9 in., at which point no *S* was able to traverse the distance without jumping. Throughout this period, the following system of forced and free jumps was adopted to counteract positional bias: (1) free jump; (2) forced jump to non-preferred side (on jump 1); (3) forced jump to preferred side (on jump 1); (4) free jump; (5) forced jump to preferred side (on jump 4); (6) forced jump to non-preferred side (on jump 4). This systematic procedure seemed to have accomplished its purpose. Almost all the *Ss* showed considerable flexibility with respect to positional preferences. On the

TABLE II

MEDIAN PERFORMANCE OF THE FOUR GROUPS

	Group I	Group II	Diff.	Group III	Group IV†	Diff.
Initial errors	29	42	13*	143	269	126†
Total errors	48	74	26*	185	338	153†
Days	15	18	3	56	96	46†

* Significant beyond the 5% level (Festinger's test).

† Significant beyond the 1% level (Festinger's test).

‡ At the conclusion of the experiment, 9 of the 13 *Ss* of Group IV had failed to achieve the criterion of learning on this problem in 96 days of training. The values entered here for Group IV thus represent underestimates of those which would have obtained had training been continued. For purposes of the significance test, the *Ss* not reaching criterion were assigned maximal ranks.

day following this phase of training the *Ss* were randomly assigned to the four experimental groups. Group I ($N=14$) was trained on Problem I; Group II ($N=12$) on Problem II; Group III ($N=13$) on Problem III; and Group IV ($N=13$) on Problem IV.

Discriminative training. The *Ss* of each group were given 8 trials per day, 2 trials to each stimulus-pair, with each day's stimulus-arrangements being alternated following one of a series of modified Gellermann orders.⁹ On the day prior to the first day of discriminative training, each *S* was given 8 forced-correct jumps, 2 to each of the stimulus-pairs, for the purpose of teaching it to jump at cards. In discriminative training, a correctional method was used throughout, with both initial and total errors being recorded. On any given trial, an *S* was permitted a maximum of three incorrect responses before being guided in the correct direction. Each trial was thus terminated by a correct response. The criterion of learning was two errorless days.

Results. The course of learning in each group on each problem is plotted in Fig. 1. Examination of these curves makes it apparent that the discriminative problems presented to Groups I and II were much easier than

⁹ L. W. Gellermann, Chance orders of alternating stimuli in visual discrimination experiments, *J. Genet. Psychol.*, 42, 1933, 206-207.

those presented to Groups III and IV. Further evidence of the disparity in difficulty between the two sets of problems is found in Table II. Group I (simultaneous presentation), for example, averaged only 29 initial errors, whereas Group III (simultaneous presentation) averaged 143 initial errors.

Problems I and II were not very different in difficulty. As shown in Table II, the differences between Groups I and II, tested by Festinger's

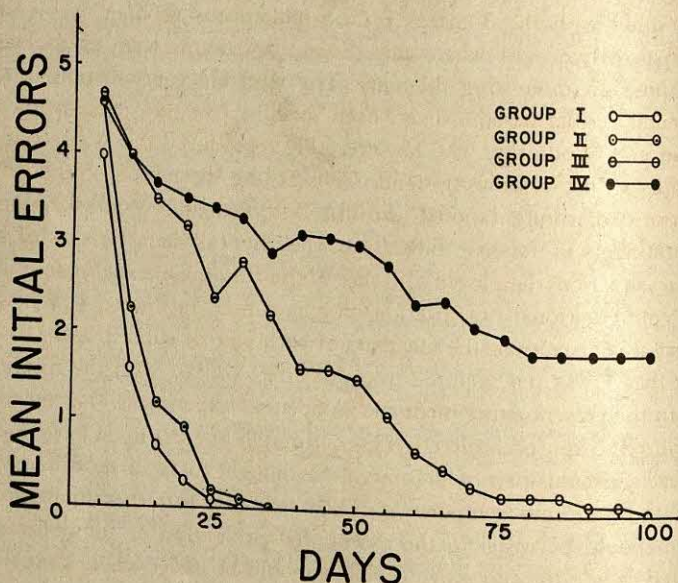


FIG. 1. THE COURSE OF LEARNING IN THE FOUR GROUPS

The curves are plotted in terms of mean values for consecutive 5-day intervals.

method,¹⁰ nowhere achieve significance much beyond the 5-% level. Problems III and IV, the more difficult problems, exhibited, on the other hand, marked differences in difficulty. The difference in performance on these two problems, measured in terms of each of the three indices, was well beyond the 1-% level.

The above comparisons are taken as evidence of the validity of the hypothesis tested; namely, that as stimulus-similarity increases, the difficulty of the successive problem increases relative to that of the simultaneous problem, while the absolute difficulty of both problems is increased.

¹⁰ Leon Festinger, The significance of difference between means without reference to the frequency distribution function, *Psychometrika*, 11, 1946, 97-105.

Additional support also is given to the finding of Saldanha and Bitterman to the effect that, with relatively similar stimuli, discrimination without opportunity for comparison is in many cases impossible. Of the 13 Ss in Group IV, 9 failed to reach the criterion in 96 days of training.

Weise and Bitterman found that under the conditions of their experiment, the performance of a simultaneous group was significantly inferior to that of a successive group. In the present experiment (as well as in Spence's study with the T-maze¹¹) the simultaneous problem was easier than the successive even when quite dissimilar stimuli were used. These findings pose an interesting dilemma. The apparent contradiction in the experimental results has, however, been given a tentative explanation by Bitterman and Wodinsky, who have recently reproduced the Weise-Bitterman results with the jumping-stand.¹² Discarding the notion of the functional priority of configurational learning, Bitterman and Wodinsky sought for determinants of relative difficulty in the physical characteristics of the particular situations employed. In the Weise-Bitterman experiment, they argued, configurational organization may have been facilitated (1) by the close spatial contiguity of the members of each pair of stimuli, and (2) by the fact that S was not required to make direct contact with the discriminanda. In the present experiment and in Spence's experiment, they noted, S was required to approach directly (jump through or enter upon) one member of each stimulus-pair, which might be thought to facilitate differentiation of the pair into components. They reasoned that the simultaneous problem should be easier to the extent that perception of components is facilitated, while the successive problem should be easier where configurational organization develops.

In a recent experiment with human Ss, Loess and Duncan obtained data in accord with the present findings on the rôle of stimulus-similarity.¹³ Their design was somewhat similar to that employed by Grice,¹⁴ *i.e.* in their simultaneous situation two stimulus-cards were presented to S, whereas in the successive situation S was presented with one stimulus-card. In both situations, however, S was required to make a positive response, whereas in Grice's experiment the successive situation demanded a positive response in one stimulus-situation and inhibition of response in the other.

¹¹ Spence, The nature of the response in discrimination learning, *op. cit.*, 89-93.

¹² Bitterman and Wodinsky, *op. cit.*, 371-376.

¹³ H. B. Loess and C. P. Duncan, Human discrimination learning with simultaneous and successive presentation of stimuli, *J. Exper. Psychol.*, 44, 1952, 215-221.

¹⁴ G. R. Grice, Visual discrimination learning with simultaneous and successive presentation of stimuli, *J. Comp. & Physiol. Psychol.*, 42, 1949, 365-373.

Thus, interpretation of the successive situation of Loess and Duncan in terms of Spence's theory necessitates, as does that of Weise and Bitterman, the assumption of stimulus-compounding. Contrary to predictions which might be made on the basis of this assumption, however, Loess and Duncan found that with an easy discrimination, the method of presentation (simultaneous or successive) made no difference. With a more difficult discrimination, however, rate of learning with simultaneous presentation was significantly superior to that with successive presentation. The only difference between the two sets of stimulus-cards for the easy and for the more difficult problems was the addition of triangles (half with apex up and half with apex down) to each of the cards for the more difficult problem. It is perhaps self-evident that this change sufficed to make the cards for the more difficult problem more similar to one another, thus accounting for the greater difficulty of both simultaneous and successive problems and for the increased disparity in difficulty between the two problems.

SUMMARY

Four groups of animals were given discriminative training on the jumping apparatus. Group I (simultaneous presentation) and Group II (successive presentation) learned to discriminate between relatively dissimilar stimuli—horizontal versus vertical stripes and large versus small circles. Performance of Group I was somewhat superior to that of Group II, although the significance of difference in terms of trials and errors fell short of the 1-% level of confidence. Group III (simultaneous presentation) and Group IV (successive presentation) were trained with relatively similar stimuli—light versus dark gray and thick versus thin stripes. Group III achieved the learning criterion while Group IV did not. The results support the hypothesis that as stimulus-similarity increases, the difficulty of the successive problem increases relative to that of the simultaneous problem, while the absolute difficulty of both problems is increased. Since opportunity for direct comparison of the stimuli to be discriminated is afforded by simultaneous presentation but not by successive, this experiment provides further support for the concept of relational learning in the rat.

VISUAL RECOGNITION OF VERBALLY LABELED FIGURES

By W. C. H. PRENTICE, Swarthmore College

Some twenty years ago, Carmichael, Hogan, and Walter published a study on memory for figures which was especially interesting because it threw light on a controversy that was started by Wulf's report of systematic changes in memory traces.¹ Carmichael and his associates showed simple line drawings to their Ss with a spoken name for each drawing, each of two groups being given a different name for the same figures.² When S was then asked to draw what he had seen, comparison proved that the reproductions had been selectively influenced by the different labels. This result made plausible the suggestion, if it did not confirm it, that assimilation of one memory to another, or to an archetype, might be a fairly simple associative matter.

In the intervening years since then, the nature of Wulf's problem has been clarified by the work of Hanawalt, Zangwill, and Hebb and Foord.³ We now know that whatever changes take place under the influence of conditions like those of Wulf, they operate on S's response in the testing-situation (usually the drawing of a picture) rather than on the process of memory that makes such reproductions possible. This conclusion is based on the finding that the process of recognition is unaffected by the very influences that make such dramatic changes in reproduction.

The assimilation to verbal symbols, demonstrated by Carmichael, Hogan, and Walter, has seemed to some authors, however, to be a different kind of change. Graham, for example, implies that the results of Carmichael *et al.* represent a perceptual change.⁴ Such an interpretation would create real difficulties for perceptual theory if it were sustained, but the

* Accepted for publication March 25, 1953.

¹ Friedrich Wulf, Über die Veränderung von Vorstellungen (Gedächtnis und Gestalt), *Psychol. Forsch.*, 1, 1922, 333-373.

² L. C. Carmichael, H. P. Hogan, and A. A. Walter, An experimental study of the effect of language on the reproduction of visually perceived form, *J. Exper. Psychol.*, 15, 1932, 73-86.

³ N. G. Hanawalt, Memory trace for figures in recall and recognition, *Arch. Psychol.*, 31, 1937 (No. 216), 1-89; O. L. Zangwill, A study of the significance of attitudes in recognition. *Brit. J. Psychol.*, 28, 1938, 12-17; D. O. Hebb and Esme Foord, Errors of visual recognition and the nature of the trace, *J. Exper. Psychol.*, 35, 1945, 335-348.

⁴ C. H. Graham, Visual perception, in S. S. Stevens, *Handbook of Experimental Psychology*, 1951, 868-916, esp. 912.

specific evidence on recognition referred to above makes it immediately suspect. If the alleged autonomous changes of Wulf can be demonstrated only by the method of reproduction, it seems not unlikely that the same is true of assimilative effects. The experiment of Hanawalt and Demarest corroborates this interpretation.⁵ They showed that if Carmichael, Hogan, and Walter's figures are exposed without labels, the directional distortions nevertheless occur in the reproductions, if *E* asked at the time of recall: "Now draw the one that looked like [for example] a canoe." This result shows that it is not necessary for *S*'s original percept to be distorted. It is thus legitimate to inquire whether the labeling used by Carmichael *et al.* does in fact operate at the time of original observation or only later. Their results could be due only to specific influences on the act of reproduction, in which case a test by recognition should not be expected to show any more errors in one direction than in another.

In the present experiments, the recognition-technique has been used with a procedure that is otherwise very much like that of Carmichael *et al.* The purpose was to discover whether the systematic distortions produced by naming may be absent when *S* does not have to draw the figures.

Procedure. The same figures used by Carmichael and his associates were used in this investigation. They were redrawn from his article, approximately 2 in. high on 4-in. squares of white cardboard, and presented one at a time to *S* who sat across a table from *E*. Each card was exposed by hand for 4 sec., and the entire series was repeated twice in the same order.

Instructions. The instructions were as follows: "I am going to show you some simple line drawings. I wish you to look at each one with the idea of being able to recognize it when you see it again. As I present each one, I will tell you something about what the next one will look like. The first one looks something like a *bottle (stirrup)*." While the first picture was exposed, *E* was preparing to turn up the second one and announcing, "The next one is something like a *canoe (kidney bean)*." After the entire set had been shown, *E* said, "Now I will go through them once more that you will have a better chance to remember them." The second trial was identical with the first.

Immediately at the end of the second run, *E* said, "Now I am going to show you a set of figures that includes the ones you have just seen. I want you to pick out the ones you saw and read their numbers off to me." *E* then exposed on the table two large sheets containing 60 drawings, each in a 4-in. square and drawn to the same scale as those in the original series. The set of 60 represented one each of the original 12 figures plus four other examples of the same general type: two drawn to approximate in varying degree the label used respectively for each experimental group. Thus there were five examples of the figure that was sometimes called a 'hat' and sometimes called a 'beehive.' Two of these were so drawn as to

⁵ Hanawalt and I. H. Demarest, The effect of verbal suggestion in the recall period upon the reproduction of visually perceived forms, *J. Exper. Psychol.*, 25, 1939, 159-174.

exaggerate those features that would make the picture look like a 'beehive,' and two were drawn to obscure those features and emphasize the similarity to a 'hat.' In some cases, the extreme types were borrowed from Carmichael and his associates' examples of the distortions drawn by their Ss. In other cases *E* had to invent what seemed to be likely confusions. The figures were arranged in orderly rows, but they were unsystematic with respect to any of the variables of the experiment, some effort being made to 'randomize' the order. Every figure was numbered according to its position to simplify recording.

S was allowed as much time as he wished. The task of selecting 12 out of 60 figures typically occupied about 3 min., with the range being between 1.5 min. *E* recorded the number of every figure chosen by *S*, keeping track of any changes explicitly announced.

In most cases, *S* chose exactly 12 figures, one from each set. In a few cases, he chose less than 12. In a somewhat larger number of cases, he chose more than one form for one or more of the varieties. (Total choices for 60 Ss was 890, or about

TABLE I
NAMES ASSIGNED TO FIGURES (IN ORDER OF PRESENTATION) FOR TWO
DIFFERENT GROUPS OF Ss

List 1	List 2
stirrup	bottle
canoe	kidney bean
hat	beehive
table	hour glass
diamond in a rectangle	curtains in a window
dumbbell	eyeglasses
trowel	pine tree
eight	two
broom	gun
four	seven
letter C	crescent moon
sun	ship's wheel

15 per *S*.) The last example chosen of any one variety was accepted as the one to be counted except in a few cases where *S* clearly misunderstood the instructions and gave several examples of each variety. In those cases, the instructions were repeated with emphasis on the fact that the group contained exact copies of the original figures. *S* then went back and chose the ones he thought exactly like the originals, and his choices from that point on are included in the main tabulation. His earlier choices are included with the 'tentative' choices (see below). Data from 6 Ss making this error are included.

All Ss were drawn from an elementary class in psychology at the University of California; all were naïve with respect to the problem; none had seen the figures before. Of the 60 Ss, half (selected by lot) were given the labels in List 1 and half those in List 2. The lists are presented in Table I. To test a hypothesis about the subjective nature of the task (see below), *E* gave half the Ss of each group the following special instruction as they began the recognition-test: "Don't bother about the names. They were only to help you learn the pictures. Just pick out the ones you saw."

Results. Because the design of the figures necessarily showed considerable arbitrariness, it was anticipated that Ss might be able to choose the correct

ones so easily as to provide no test of the rôle of the labels. Fortunately, errors were frequent, and it was not necessary to modify the original group of test-figures. No *S* chose all 12 correctly, and the group as a whole averaged 3.7 errors. On one figure, there were almost no errors (the table-hourglass figure) perhaps because it was so simple and symmetrical in its original form. On two figures the tendency to choose examples that varied in one direction from the original seems significantly greater than that in the other direction. One such would favor 'assimilation' for List 1; the other for List 2. Otherwise, the tendency to choose erroneous drawings seems to be equal for the examples drawn to approximate the two labels.

The errors were scored as 'positive' or 'negative' depending on whether they deviated from the original in the direction of the label given that particular *S* or in the other direction. The results are displayed in Table II.

TABLE II
CORRECT CHOICES AND TWO TYPES OF ERROR IN RECOGNITION
($N=60$)

Correct choices	457
Choices of figure more like label	121
Choices of figure less like label	100
Total	678*
Mean 'positive' choices	2.01
Mean 'negative' choices	1.67

* The difference between this number (678) and 12 times the number of *Ss* (720) results from the omission by some *Ss* of all examples of one or more varieties of figure. Altogether 42 such omissions were recorded, some *Ss* omitting two or three apiece.

Since many *Ss* chose more than one sample of some varieties, and the convention of accepting only the last chosen was somewhat arbitrary, a separate tabulation of the 'tentative' choices was made. This tabulation includes all figures mentioned by any *S* and not included in Table I. Some of these choices had been specifically abandoned by *S* who might say something like, "I guess I made a mistake when I took 9; may I change it to 42?" Other choices were merely superseded, in the sense that *S* would select both 9 and 42 with no apparent recognition of any conflict. As noted above, a few *Ss* each chose 25 to 40 drawings through some misunderstanding of the instructions. All of their choices are called 'tentative' except those explicitly named as exact reproductions after the instructions had been repeated. Table III shows the results of the 'tentative' choices.

As previously noted, special instructions were given to approximately half the group. One possible interpretation of the results of Carmichael, Hogan, and Walter is that *Ss* thought they were meant to draw a picture like the label rather than like the specific example they had seen. This

possibility is emphasized by the ease with which Hanawalt and Demarest reproduced the results by an instruction at the time of reproduction. In the case of recognition, *S* could take an attitude that would lead him to interpret the label as an instruction to treat the symbol like the label, and, in the test, to find the prototype rather than the original. It was hoped that a warning to disregard the label at the time of recognition might counteract any such self-instruction. As it turned out, *Ss* given the special warning behaved just like those not so warned. Table IV shows the two sub-groups.

TABLE III

CORRECT CHOICES AND TWO TYPES OF ERROR AMONG 'TENTATIVE' RECOGNITIONS

Correct choices	63
Choices of figure more like label	74
Choices of figure less like label	75
Total	212

(In this case $N = 26$ instead of 30 because an error on the part of *E* resulted in the omission of the special warning for three *Ss* who should have had it, and, in order to balance the two sets of labels for this comparison, Table IV presents data for the first 26 warned *Ss* instead of all 27.)

Several kinds of statistical treatment of these data might be considered appropriate to our problem. Three have been applied.

First, errors in the 'positive' direction have been compared to errors in the 'negative' direction for each set of drawings. Omitting two sets of

TABLE IV

CORRECT CHOICES AND TWO TYPES OF ERROR FOR 'WARNED' *Ss*

(N=26)

Correct choices	183
Choices of figure more like label	54
Choices of figure less like label	45
Total	282
Mean 'positive' choices per <i>S</i>	2.08
Mean 'negative' choices per <i>S</i>	1.73

drawings (bean-canoe and table-hourglass) each of which gave a total of less than 10 errors, a Chi-square test was applied to the hypothesis that positive and negative errors, as different in number as these, might be produced by sampling from the same population. When one-half the total number of errors for that case is used as the theoretical expectation for 'positives' in each of the 10 cases, X^2 equals 6.52 with 9 *df*. *P* is thus greater than 0.50, and there is no ground for rejecting the null hypothesis.

A second form of test is suggested by the distribution of total errors between the groups of *Ss* given the two different lists of labels. For the group given List 1, the total number of errors was 111; for the group

given List 2, it was 110. Perhaps these errors are differentially distributed if we look at individual pictures. Each erroneous picture attracting as many as 10 choices (thus permitting a theoretical value of 5 or more in each cell) was examined to see whether it was selectively chosen by Ss given the two sets of labels. There were 10 incorrect drawings that were chosen by at least 10 Ss. A Chi-square test here lumps together positive and negative cases, and in fact 6 of the 10 cases are positive and 4 are negative. Even so, X^2 turns out to be 6.74 for 9 *df*. *P* is again greater than 0.50, and there is no reason to reject the hypothesis that Ss receiving different labels make the same kind of errors.

The Chi-square tests omit some of the data. Certain drawings that attract a low percentage of choices would achieve our criterion for inclusion in the X^2 -test if we used a larger population. One way of including all those drawings that were chosen by only one or two Ss is to divide all the errors into positive and negative groups and see what the probability may be that such a distribution might result if no selective factor were at work. This may be done by using the standard error of a proportion and applying it to the raw data of Table II; in this case $N = 221$, and the hypothetical probabilities of an error being positive or negative are equal. The $\sigma_{\text{prop.}}$ is slightly less than 0.034, and the Z-score we are concerned with is thus 1.43. *P* (one-tailed) is about 0.075, above usual standards of significance despite the obvious inflations of *N*. A two-tailed test, which is probably better justified, gives the discrepancy even less significance. It is equally clear that this kind of test shows the data of Tables III and IV to be insignificant, since the deviations are smaller and *N* also is smaller.

The mean number of errors per *S* in the positive direction is almost exactly 2.00; in the negative direction, 1.67. The σ s for a *t*-test were not calculated, as the insignificance of the differences is clear from inspection.

Summary. Even with a visual recognition-test sufficiently sensitive to produce half as many errors as correct choices, using the figures and labels of Carmichael, Hogan, and Walter, it has not been possible to show that those errors are distributed in a way favoring selection of a shape more like the label used during learning. We must conclude tentatively that the effect reported by Carmichael is specific to the task of reproduction. The present evidence, combined with that of Hanawalt and Demarest, suggests that the use of labels during learning does not modify the original visual experience or its memorial process. It is more likely that labels produce a selective set affecting the activity of drawing, much as special instructions would operate.

THE USE OF COMPARATIVE RATING SCALES FOR THE EVALUATION OF PSYCHOPHYSICAL DATA

By HARRY HELSON, University of Texas, WALTER C. MICHELS, Bryn Mawr College, and ARTIE STURGEON, University of Texas

It has been shown in previous studies that the absolute type of judgment commonly employed in rating scales may be utilized for psychophysical measurements as well as in the evaluation of responses which have no physically measureable stimuli.¹ In the absolute method of judgment, although background stimuli may be introduced to serve as anchors, no direct comparison of variable and standard is required. In such cases it has been shown that the basis of judgment is the *adaptation-level*, defined as a weighted geometric mean of present and past stimuli. The question arises: Can the quantitative theory found adequate for absolute judgments be extended to direct comparison of standard and variables as in the method of constant stimulus differences? If this can be done then the existing theory of absolute judgment may be applied to the comparative type of judgment. This extension obviously requires the introduction of a new type of scale, the *comparative rating scale*, wherein judgments are in terms of comparative categories (much heavier, a little heavier, equal, a little lighter, much lighter) rather than the absolute type previously in question. If the theory of absolute judgment proves adequate to the comparative type of judgment then not only have we succeeded in extending the use of the comparative type of rating to psychophysical measurements but the way is opened up for quantitative treatment of a number of psychophysical problems, *e.g.*, time-order effects, which have hitherto stood as isolated phenomena.

Procedure. Twelve Os were required to judge five weights ranging from 300 to 500 grm. in 50-grm. steps (V) following a standard (S) of 400 or 350 grm. ($S;V_2$ time-order) or one of the standards following the five series weights (V_1S_2 time-order) according to the following categories which may be said to comprise a comparative rating scale: very much heavier; much heavier; heavier; a little heavier; equal; a little lighter; lighter; much lighter; very much lighter. For purposes

* Accepted for publication April 20, 1953.

¹ Harry Helson, *Adaptation-level as a basis for a quantitative theory of frames of reference*, *Psychol. Rev.*, 55, 1948, 297-313; W. C. Michels and Helson, *A reformulation of the Fechner law in terms of adaptation-level applied to rating-scale data*, this JOURNAL, 62, 1949, 355-368; M. C. Nash, *An experimental test of the Michels-Helson theory of judgment*, this JOURNAL, 63, 1950, 214-220.

of quantification, numbers from 1 to 9 were assigned to the response-categories, with 1 for very much lighter to 9 for very much heavier. While only the S_1V_2 results are considered in this study, the order of presentation of stimuli should be mentioned here. The 12 O s were randomly divided into two groups, A and B. Group A judged S_1V_2 followed by V_1S_2 with the 400-grm. standard and Group B judged V_1S_2 first followed by S_1V_2 . Judgments were next made with the 350-grm. standard and the two groups were reversed so that Group A judged V_1S_2 followed by S_1V_2 and Group B began with S_1V_2 followed by V_1S_2 . Averages for like conditions constitute the data of this study and are based, therefore, upon 60 observations (12 O s x 5 judgments each) of each stimulus.

The stimulus-weights were obtained by filling 6-oz. medicine bottles with sand or shot and were wrapped in paper to obliterate any visual cues as to heaviness—a precaution which may have been unnecessary since the O s were blindfolded during the observations. Stimuli such as these have been found to yield much more reliable data than the usual cylindrical weights used in weight-lifting experiments, perhaps because O s can grasp the narrow necks of the bottles easily.

A detailed set of instructions was read to each O before the start of observations, explaining how the stimuli were to be lifted and returned to position. Two series of practice trials were given for each condition after which five series were given and recorded. The instructions read to the O s were as follows: "You will be presented with two weights in succession. After lifting the second weight you are to give a judgment by calling it very much heavier than . . . equal to . . . or very much lighter than the preceding weight." The full list of categories was read twice to the O s to insure familiarity and the practice series showed that such a scale does not offer any difficulties to the O s (all graduate or senior students in psychology).

Results. The data in Table I are the averages of the comparative judgments translated into a numerical scale in the manner described above and

TABLE I
AVERAGE COMPARATIVE SCALE-RATINGS OF WEIGHT STIMULI BY

Standard (gm.)	12 O s WITH TWO STANDARDS Stimuli (in gm.)					AL (PSE)
	300	350	400	450	500	
400	3.3	4.3	5.3	6.3	7.1	381
SD_m^*	0.38	0.14	0.10	0.16	0.16	
350	4.2	5.3	6.1	7.0	7.7	338
SD_m^*	0.11	0.06	0.14	0.16	0.16	

* The values of SD_m in this table were obtained by using the averages of each of the 12 O s. The inter-observer variability is obtained if these values are multiplied by 3.46.

they show a monotonically increasing heaviness judgment as the weight of the stimulus increases from 300 to 500 gm. In this respect the results obtained by the comparative rating scale technique are formally similar to the classical method of constant stimulus differences. The two methods differ, however, in that the classical method is designed for and usually employs only three comparative judgments, heavier, equal, and lighter, and

the quantitative data represent the percentage of each category as a function of the stimuli. In our method the quantitative data refer directly to the originally given categories of judgment. The values of stimulus eliciting any one of the categories in the rating scale may be determined from our data. Hence the comparative rating-scale method yields a much wider range and variety of judgments than does the classical method and should be correspondingly more useful. The standard errors of the mean values in Table I give clear evidence of the high reliability of the comparative rating-scale method.

A check on the sensitivity of the comparative rating-scale procedure employed here is the fact that the data show the classical negative time-order error found with method of constant stimulus differences. It will be noted that the stimulus judged equal to the standard 381 grm. when the standard is 400 grm. which is the usual negative time-order effect with S_1V_2 . Similarly, adaptation-level or point of subjective equality (*PSE*) is 338 grm. when the standard is 350 grm., as is expected on the basis of the theory underlying this study. Since the time-order error is generally regarded as a small, though non-negligible, effect in measuring thresholds and *PSE*, its clear presence in the data obtained by the comparative rating method shows that this method is capable of yielding as fine differences in sensitivity as the longer, more difficult classical methods for obtaining comparable data. Moreover, the method of comparative ratings makes possible a theory of *PSE*, time-order effects, and other psychophysical phenomena which classical methods have not provided.

Theory. In reformulating the Fechner law, Michels and Helson have shown that quantitative predictions may be made of the ratings of a series of stimuli, provided only that certain general conditions are met by the rating scale itself.² Essentially these conditions are: (1) the scale must contain an odd number of categories; (2) the central category must represent a 'neutral' judgment, and the scale must be symmetrical about this category; and (3) the judgment-scale and the stimuli encountered must be *equivalent* in the sense that the scale be broad enough to include judgments of all stimuli encountered and yet be so narrow that its extreme values do not fall outside the range of judgments elicited by any of the stimuli. Nash has shown that a scale, such as: extremely light; very light; medium light; medium; medium heavy; heavy; very heavy; extremely heavy; satisfies these conditions for the single stimulus method in weight judgments.³

² Michels and Helson, *op. cit.*, 360.

³ Nash, *op. cit.*, 219.

An implicit assumption in the earlier treatment is that the judgments of single stimuli are made in terms of the adaptation-level as a standard. At first glance, it might appear that, when comparative judgments are made involving a standard and a set of variables, the standard would merely replace the adaptation-level as the basis of comparison. Such an expectation would require, however, the further assumption that the person forming the judgment has been completely successful in suppressing all adaptation-effects except those resulting from the standard stimulus. Since adaptation is known to play a highly significant rôle in all perception, it would be very surprising if this complete success were achieved. It seems more reasonable to suppose that in the comparative judgment neither the adaptation-level (AL) found in single stimulus judgments (SAL) nor the standard (S) will be the effective AL and evoke the equality judgment, but rather that a new comparative adaptation-level (CAL) will be established. We may expect the CAL to lie between SAL and S , its exact location depending on the relative weights of SAL and S in establishing the new AL .

Quantitative treatment. It has been shown that judgments of a series of stimuli, assuming a properly constructed rating scale, can be represented by the relation:⁴

$$J = C + K' \log (X/A'), \dots \dots \dots [1]$$

where J is the judgment value on a rating scale running from 1 to $1 + 2N$, $C = 1 + N$, X is the magnitude of the stimulus, A' is the series-adaptation-level (SAL), and K' is a determinable function of N and of the conditions under which the judgments are made. Equation [1] has been tested under a variety of conditions and has been found to be valid to a fairly high degree of accuracy.⁵

In accordance with the discussion of the last section, we must replace A' by the comparison adaptation-level (CAL) if a comparative scale is used. The relation of CAL , S , and SAL is indicated by the fact that the adaptation-level is found to be a weighted geometric mean of all the stimuli, past or present, which affect the judgment. Hence we can probably take CAL as a weighted geometric mean of A' and S , i.e.:

$$(CAL) = A_1 = S^s A'^t, \dots \dots \dots [2]$$

where

$$s + t = 1 \dots \dots \dots [3]$$

The exponent s is then a measure of the influence of the standard stimulus, t is a measure of the influence of the SAL . If the judgments are completely in terms of the standard we shall have $s = 1$, $t = 0$; if the standard has no effect, we shall have $s = 0$, $t = 1$. In any practical case, we may expect values between these limits.

We may now rewrite Equation [1] in a form applicable to comparative judgments. We first express the quality of Equation [2] in logarithmic form:

$$\begin{aligned} \log A_1 &= s \log S + t \log A' = s \log S + (1-s) \log A' \\ &= s \log (S/A') + \log A' \dots \dots \dots [2A] \end{aligned}$$

⁴ Michels and Helson, *op. cit.*, 361.

⁵ Michels and Helson, *op. cit.*, 361-366; Nash, *op. cit.*, 216-218.

Substitution of A_1 into Equation [1], in place of A' gives:

$$J = C + K' \log (X/A_1) \dots \dots \dots [4]$$

$$= C + K' \log (X/A') - K's \log (S/A'), \dots \dots \dots [4A]$$

as the prediction of the comparative judgment J as a function of the stimulus, the standard, and the series adaptation-level. Equations [4] and [4A] do not give a complete determination, however, since A' is not the same as the adaptation-level of the series *without a standard*. Hence the only prediction that we can make without

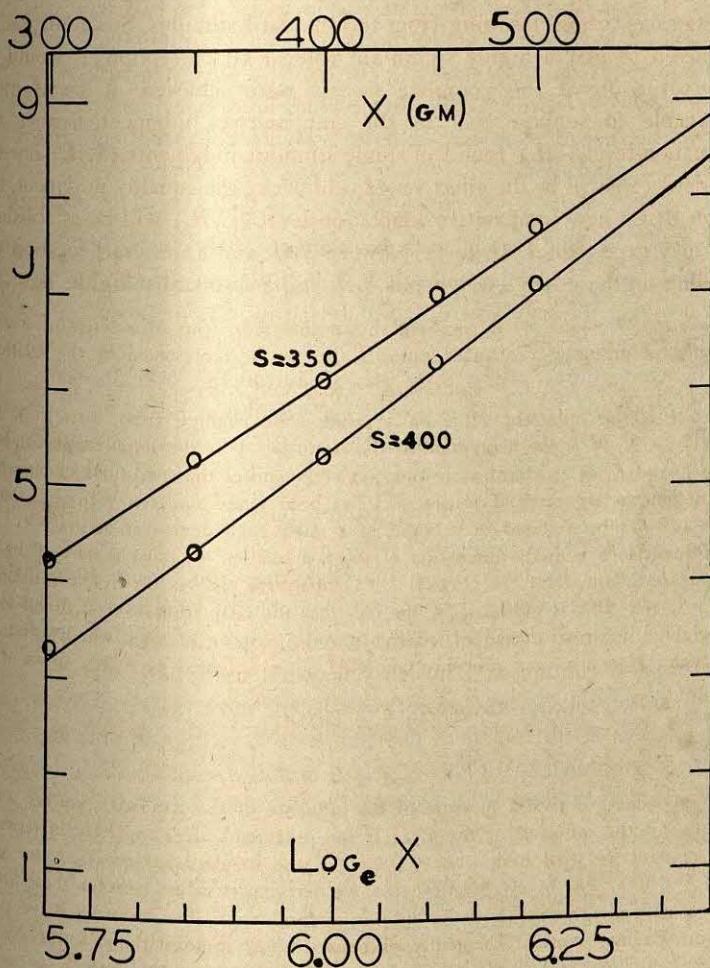


FIG. 1. JUDGMENTS OF FIVE STIMULI BY COMPARISON WITH 350 GRM. AND 400-GRM. STANDARDS

The solid lines are least square fits to the data, and follow the relations: $J = 6.6 \log_e X - 33.4$ for the 350-grm. standard; $J = 7.5 \log_e X - 39.6$ for the 400-grm. standard.

further information is that the curve of comparative judgments against $\log X$ should be parallel to the corresponding curve obtained for the series by the method of single stimuli and therefore parallel to each other when different standards are used.

Equations [4] and [4A] have some interesting consequences in regard to time-order errors. These will be discussed in a later paper, which will also develop a method for the determination of the weighting factors s and t .

Fig. 1 shows a plot of the data obtained with the two standards. The lines shown have been fitted to the experimental points by the method of least squares. Their slopes are 7.5 for the 400-grm. standard and 6.6 for the 350-grm. standard.⁶ While these values are not in perfect agreement, they do not depart from each other to a significant extent, and the curves are close to parallelism.

Conclusions. The present study has demonstrated that reliable comparative rating judgments may be made of stimuli when a standard is provided, thus bridging the gap between absolute rating-scale methods and classical psychophysical methods.

The comparative rating scale has an advantage over the method of absolute judgment because it provides a stimulus-standard which helps to stabilize the adaptation-level and therefore the judgments of stimuli. This advantage is not as great as might have been expected, due to the fact that series stimuli and residual or background factors, as well as the standard, affect the comparative adaptation-level.

The classical methods of liminal determination, while they provide methods for determining thresholds or number of *jnds* above or below a standard, do not establish equivalence of stimuli and response in a psychological continuum. Thus they may tell us by how many *jnds* two stimuli differ, but no one, except perhaps the most practiced *O*, can make even a guess at this number. The comparative rating-scale method makes possible determination of the values of stimulus which will be judged to differ from a standard by specifiable psychological intervals.

⁶ The fact that the average slope is 7.0 is itself of some interest. According to the theory of absolute judgments, the maximum slope for a nine category scale should be 3.5, just about half of that found in the present study. The reason for the disagreement is that the equivalence condition of assumption (3) above was not met in this experiment, the range of stimuli having been limited to approximately that used for comparisons by classical methods. Reference 2 (p. 365) predicts that abnormal slopes will occur with a narrow range of stimuli, and contains data, obtained under conditions similar to those of this study, which yielded a slope of 6.9.

A QUANTITATIVE THEORY OF TIME-ORDER EFFECTS

By WALTER C. MICHELS, Bryn Mawr College, and
HARRY HELSON, University of Texas

The time-order effect (*TOE*), commonly referred to as time-order error, has been the subject of much investigation and has been extensively discussed in the literature.¹ It has both practical and theoretical importance. One of the most valuable practical results of work on *TOE* has been the knowledge that comparisons must be made in two time-orders if errors are to be minimized. In the absence of a quantitative theory, it has been customary to assume, however, that the *TOE* for standard first, variable second (S_1V_2) will be exactly equal to *TOE* for variable first, standard second (V_1S_2).² If this assumption is incorrect, serious errors may creep into psychophysical measurements designed to establish psychological equality between physically different stimuli; for example, in the establishment of the luminosity function of the C.I.E. standard observer.³ The theoretical importance of *TOE* lies in its implications for psychological scaling and for theories of perception and judgment.

The successful application of rating-scale techniques and adaptation-level (*AL*) theory to comparative judgments makes possible the development of a treatment of *TOE* which is quantitative and which integrates *TOE* with other psychological phenomena.⁴ The present study was undertaken to test the validity of an extension of adaptation-level theory to time-order effects.

Experimental test. The detailed procedure by which the data of this study were obtained has been described in an earlier paper.⁵ Lifted weight data were obtained

¹ R. S. Woodworth (*Experimental Psychology*, 1938, 438-449) has mentioned the most important studies up to 1938. In this study we shall limit ourselves to discussion and treatment of time-order effects rather than of the temporal dependence of these effects, which may be called the time-error.

² A less detailed quantitative theory of *TOE* than the one to be presented here has been proposed by one of us in an earlier publication: Harry Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29.

³ Similar considerations apply to space-errors where the assumption of symmetry may be equally unjustifiable.

⁴ Helson, W. C. Michels, and Artie Sturgeon, The use of comparative rating scales for the evaluation of psychophysical data, this JOURNAL, 67, 1954, 324-326.

⁵ Idem.

with 12 *O*s who judged a series of five weights ranging from 300 to 500 gm. in 50-gm. steps. Two standards of 400 and 350 gm. each were employed in two time orders, S_1V_2 and V_1S_2 . The 12 *O*s were divided into two groups to provide a balanced presentation with the two time-orders. The weights were judged in terms of a comparative rating scale containing the following categories: very much heavier; much heavier; heavier; a little heavier; equal; a little lighter; lighter; much lighter; very much lighter. For purposes of computation these categories were translated into a numerical scale from 1 to 9; 1 for very much lighter and 9 for very much heavier, with the intermediate categories numbered accordingly. The

TABLE I
AVERAGE COMPARATIVE SCALE RATINGS OF STIMULUS-WEIGHTS BY 12 *O*s
WITH TWO STANDARDS

Standard (in grm.)	Order	Stimulus-Weights (in Grm.)					AL (PSE)	Slope	
		300	350	400	450	500			
350	S ₁ V ₂	J	4.2	5.3	6.1	7.0	7.7	338	6.6
		σ _m [*]	0.11	0.06	0.14	0.16	0.16		
	V ₁ S ₂	J	6.0	5.3	4.6	3.8	3.1	365	-5.5
		σ _m	0.13	0.07	0.10	0.20	0.09		
400	S ₁ V ₂	J	3.3	4.3	5.3	6.3	7.1	381	7.5
		σ _m	0.38	0.14	0.10	0.16	0.16		
	V ₁ S ₂	J	6.5	5.9	5.4	4.8	3.9	417	-5.0
		σ _m	0.14	0.15	0.14	0.06	0.10		

* The values of σ_m in this table were obtained by using the averages of each of the 12 *O*s. The inter-*O* variability is obtained if these values are multiplied by 3.46.

data in Table I are averages of the average judgments of 12 *O*s judging each of the stimuli five times and are thus based upon 60 observations.

The data in Table I show that the *TOE* found with classical method of constant stimulus differences in lifted weights emerges clearly when judgments are made in terms of a comparative rating scale. They indicate that *O*s are able to use such rating scales as interval scales with great reliability, as is shown by the small values of σ_m .

In order that a consistent convention as to the sign and meaning of *TOE* be followed, it will be convenient to define *TOE* for S_1V_2 as:

$$E_1 = A_1 - S;$$

and for V_1S_2 as

$$E_2 = S - A_2,$$

where the subscripts refer to first or second presentation of the standard and where A_1 and A_2 are the *AL* or *PSE* in the two time-orders.⁶ Reference to the data of Table I shows that in the case of both 350 and 400 gm. as standards *TOE* is negative in the two time orders. If A_1 and A_2 for these data are averaged the resultant average is 399 for the 400-gm. standard and 351.5 for the 350-gm. standard.

⁶ Helson, Michels, and Sturgeon, *op. cit.* 321-326.

The close agreement between the averages of A_1 and A_2 and the standards must be considered fortuitous. We shall show later that such agreement is not to be expected in general. In anticipation of the quantitative treatment, we may see a reason for this by inspection of the data in Table I. The average difference between the judgments of adjacent stimuli is distinctly less in the case of V_1S_2 than it is in the case of S_1V_2 , for both standards. This means that a greater change in stimulus is needed in the V_1S_2 order than in the S_1V_2 order to bring about a specified change in judgment. This asymmetry causes A_2 to differ more from S than does A_1 .

Quantitative theory. It has been shown that the quantitative adaptation-level theory previously developed for absolute judgments of single stimuli,⁷ can be extended to comparative judgments.⁸ In particular, it was demonstrated that the judgment J , expressed as the ordinal number in a progressive rating scale which is symmetrical about 'equal,' can be expressed as a logarithmic function of the stimulus X as follows:

$$J_1 = C + K' \log (X/A_1). \dots\dots\dots [1]$$

Here C has the numerical value assigned to the equality judgment, K' is a constant, and A_1 is the *comparative adaptation-level (CAL)*, which is the *AL* when the members of a series are compared with a standard. The value of A_1 depends both on the series adaptation-level A' and on the standard S with which the stimulus X is being compared.

It will be noticed that Equation [1] predicts that the equality judgment will be aroused by a stimulus equal to the *CAL*, rather than by a stimulus equal to the standard. Hence the point of subjective equality is higher than the standard if *CAL* is more than S , and is lower if *CAL* is less than S . It has also been shown⁹ that

$$CAL = A_1 = S^s A'^t, \dots\dots\dots [2]$$

where A' is the series adaptation-level (*SAL*) which would be found by method of so-called absolute judgment, and s and t are the relative weights of S and *SAL* in determining A_1 and are subject to the condition:

$$t = 1 - s.$$

Hence A_1 lies between *SAL* and S . We therefore expect either a positive or negative *TOE* may be found depending on whether *SAL* lies above or below S . The relative positions of *SAL* and S will depend on a number of factors, such as the adaptation-coefficients and the relative magnitudes of the residual and standard stimuli. The negative *TOE* found in this experiment and in most comparisons of weights by classical methods, for example, is accounted for by the fact that the residual stimulus for lifted weights is lower than the mean of the stimuli usually employed in the comparisons.¹⁰ Hence the *SAL* for the series of weights is likely to be less than a standard which falls at or near the center of this range.

The comparative rating scale technique discussed by Helson, Michels, and Sturgeon can be used to considerable advantage in the study of *TOE*. Suppose that a series of

⁷ Michels and Helson, A reformulation of the Fechner law in terms of adaptation-level applied to rating-scale data, this JOURNAL, 62, 1949, 355-368.

⁸ Helson, Michels, and Sturgeon, *op. cit.*, 321-326.

⁹ Idem.

¹⁰ M. C. Nash, An experimental test of the Michels-Helson theory of judgment, this JOURNAL, 63, 1950, 214-220.

n stimuli $X_i (i = 1, 2, 3, \dots, n)$ are presented for judgment in comparison with a standard S , the standard being presented immediately before each $X_i (S_1V_2)$. It has been shown¹¹ that the *CAL* is then specified by:

$$\log A_i = s \log (S/A') + \log A', \dots \dots \dots [2A]$$

and that the predicted judgments are

$$J_1 = C + K' \log (X/A') - K's \log (S/A'). \dots \dots \dots [3]$$

The point of subjective equality is clearly A_1 , as we may verify by setting $J_1 = C$ and $A_1 = X$ in Equation [3] to obtain Equation [2A].

Now suppose that a new set of judgments is made, S being presented immediately after each $X_i (V_1S_2)$ and that the judgment is of S relative to X_i . The stimulus being judged now has a constant magnitude S , but the comparison adaptation-level will be different for each X_i . The rôles of S and X_i have been interchanged, and we can predict the new judgments by interchanging the two symbols representing them in Equation [3] to obtain the predicted judgments:

$$J_2 = C + K' \log (S/A') - K's \log (X/A'). \dots \dots \dots [4]$$

To find the new *CAL* for the series, we set $J_2 = C$ and set $X = A_2$ to obtain

$$\log A_2 = (1/s) \log (S/A') + \log A'. \dots \dots \dots [2B]$$

The value of A' in Equations [3] and [4] cannot be determined directly from comparisons, nor can absolute judgments of the series be expected to yield the correct value of A' , since the *AL* of the series will be changed both by the introduction of the standard and by the emphasis placed on the standard by the instructions. Fortunately, however, a symmetry condition exists and may be used to eliminate A' . Regardless of the existence of a time-order error, a stimulus which is physically equal to the standard cannot be distinguished from the standard. It is possible that the stimulus will be judged heavier (or lighter) than the standard if presented as S_1V_2 , but then the standard must be judged heavier (or lighter) than the stimulus if presented as V_1S_2 . This statement would be true if the classical technique were used; with the present method it can be extended to state that the two judgments will be quantitatively equal to each other, though neither may be the equality judgment.

In view of these considerations, we can state unambiguously that $J_1 = J_2$ when $X = S$. In other words, the two curves of J vs. $\log X$ should intersect at a point which has an abscissa $\log S$. The consistency of the equations with this statement is shown if we replace X by S and equate J_1 and J_2 , in which case we obtain an identity.

We now find it advantageous to rewrite Equation [4] in a form analogous to Equation [1], with the help of Equation [2B]. This procedure yields:

$$J_2 = C - K's \log (X/A_2). \dots \dots \dots [5]$$

A comparison of Equations [1] and [5] brings out two interesting points. In the first place, the plots of the judgments obtained with the S_1V_2 and V_1S_2 procedures should exhibit slopes of K' and $-K's$, respectively. Hence the ratios of the observed slopes may be used to determine the weight s of the standard in determining *CAL*. Secondly, by setting $J_1 = J_2$ and replacing X by S , we find that

¹¹ Helson, Michels, and Sturgeon, *op. cit.*, 321-326.

$$\log (S/A_1) = -s \log (S/A_2),$$

or

$$s = (\log S - \log A_1) / (\log A_2 - \log S) \dots \dots \dots [6]$$

This relation gives us a second method of determining s for any particular combination of stimuli and standard.

Verification. Figs. 1 and 2 show plots of the data obtained with the 350- and 400-grm. standards respectively. The lines shown have been fitted

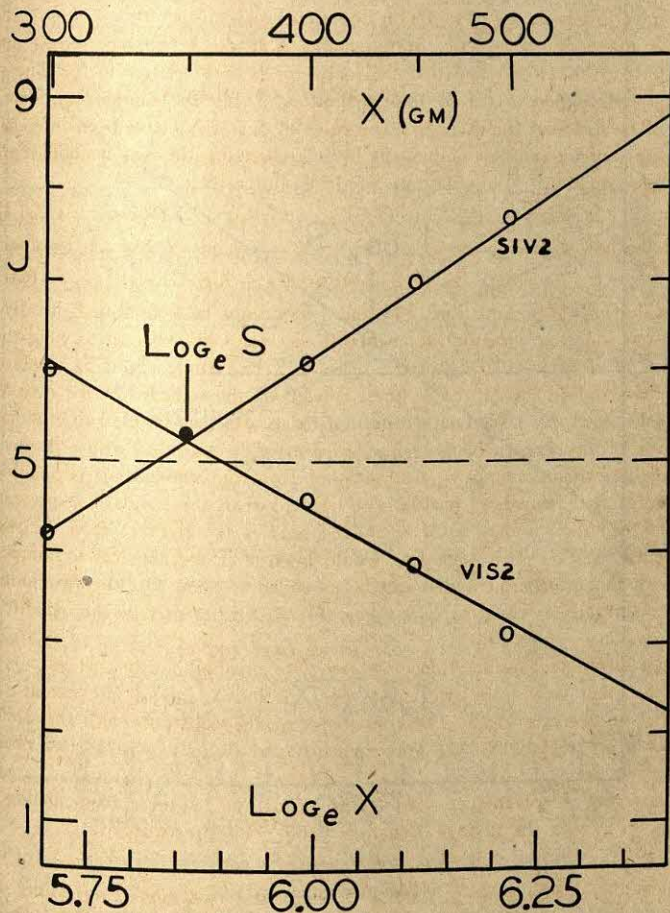


FIG. 1. COMPARATIVE JUDGMENTS OF FIVE STIMULI
(350-grm. standard)

The solid lines are least square fits to the experimental points. The average judgment shown by the black circle was identical for the two orders. The J values for S_1V_2 represent judgments of X relative to S ; the J values for V_1S_2 represent judgments of S relative to X .

to the experimental points by the method of least squares. The last two columns of Table I contain the parameters K' and AL as they have been fitted. It will be noticed that the curves in Fig. 1 for the two time-orders

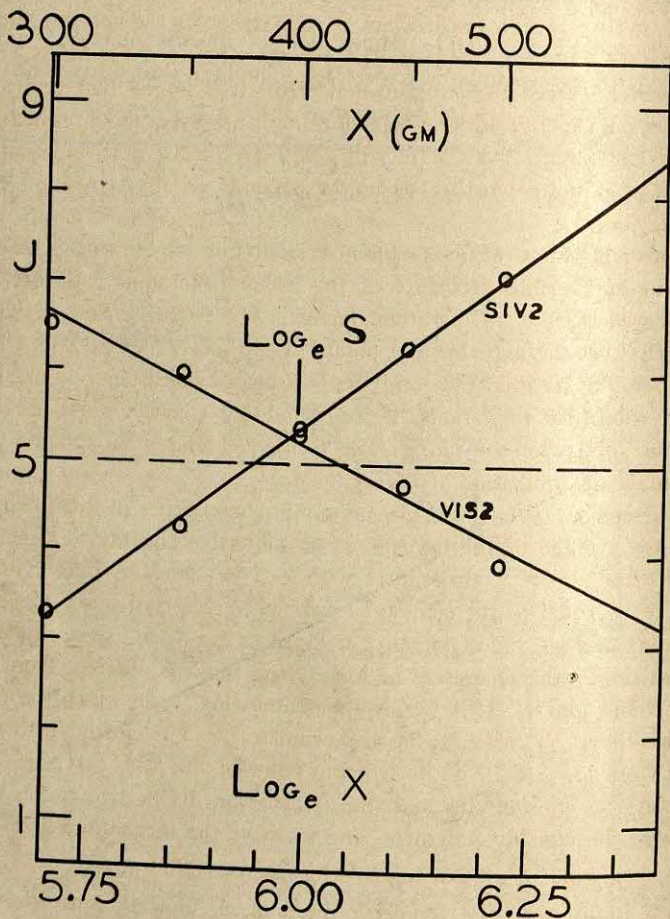


FIG. 2. COMPARATIVE JUDGMENTS OF FIVE STIMULI
(400-grm. standard)

The solid lines are least square fits to the experimental points. The J values for S_1V_2 represent judgments of X relative to S ; the J values for V_1S_2 represent judgments of S relative to X .

intersect almost exactly at $\log S$. The determination of s from the ratio of the slopes gives a value of 0.8; a second determination, based on

Equation [6], gives 0.9—a very satisfactory agreement for the 350-grm. standard. In Fig. 2 the point of intersection for the two time-orders departs from $\log S$ only by an amount within the experimental error. In Equation [6] the value of s obtained is 1.0. The ratio of the slopes gives $s = 0.7$, so that the agreement is not as satisfactory for the 400-grm. standard as for the 350-grm. standard. The difference between 1.0 and 0.7 does not, however, exceed the experimental error. It is hoped that future experiments, in which a wider range of stimuli are used and in which the standards are further from SAL , will give more precise determinations of s , and will indicate whether or not it depends on the separation of S and SAL .

The average value of s as obtained from this experiment is 0.85 ± 0.07 . The appreciable departure of this value from unity indicates that SAL , as well as S , is an important factor in the establishment of PSE .

The fact that the experimental points fall so nicely along straight lines on the semi-logarithmic plot and that reasonable consistency is obtained in the various determinations of the weighting coefficient indicates the successful application of rating scale techniques and of adaptation-level theory to the determination of TOE .

Average value of the TOE. An important consequence of this treatment is that the average of the two points of subjective equality, obtained by the S_1V_2 and the V_1S_2 procedures with a given standard, cannot be expected to be equal to the standard except in very special cases. Since the two curves of J vs. $\log X$ intersect at $\log X = \log S$, and since they have different slopes, the separation of $\log A_2$ and $\log S$ is different from that of $\log S$ and $\log A_1$. This fact is apparent in Fig. 1, in which the point at which the V_1S_2 curve crosses the ordinate $J = 5$ is clearly further removed from $\log S$ than is the crossing point of the S_1V_2 curve. If TOE is negative, as it is in this case, the intersection of the two curves must fall *above* the equality judgment, and we have the inequality relation:

$$\log A_2 - \log S > \log S - \log A_1, \text{ or } A_2/S > S/A_1.$$

Hence S is *less* than the geometric mean of A_1 and A_2 . Since the geometric mean is *always* less than the arithmetic mean, it follows, if TOE is *negative*, that

$$S < (A_1 + A_2)/2.$$

If TOE is positive, no such general conclusion can be drawn. In this case, the two curves must intersect *below* the equality judgment, so that $\log A_2 < \log S < A_1$. Because of the lower slope of the curve obtained with the V_1S_2 procedure, we find the inequality:

$$\log S - \log A_2 > \log A_1 - \log S.$$

It follows that S is *higher* than the geometric mean of A_1 and A_2 if TOE is *positive*, but we cannot say whether it falls above or below the arithmetic mean unless we have the numerical magnitudes of A_1 , A_2 , and s .

Whenever TOE is small, say less than 5% of S , both the effect of the different slopes and the difference between the geometric and arithmetic means will also be small. Under these circumstances, the mean PSE will be very close to the standard and the usual averaging procedure is justified. In all other cases, the comparative rating scale method offers the advantage that it demonstrates by how far the mean departs from the standard, and it makes possible determination of S from A_1 and A_2 by the use of Equation [6] whenever the standard cannot be measured directly.

Summary. The time-order error can be accounted for as being due to the rôle which adaptation-level plays in establishing PSE . The PSE is found to be a weighted mean of all stimuli, past or present, which affect the judgment. The standard, while exerting predominant influence, is only one of the determinative factors.

Time-order errors have been studied by the use of a comparative rating scale. When the order of presentation of the standard and the stimulus is reversed, a new relation between stimulus and judgment is found. The only judgment that remains unaffected is that of a stimulus equal to the standard. If this stimulus evokes a judgment of equality, PSE or AL is equal to the standard and there is no TOE ; if it evokes a higher judgment, TOE is negative; and if it evokes a lower judgment, TOE is positive. The average of the two PSE s obtained with the different time-orders is not generally equal to the standard, although it may be in special cases.

DISCRIMINATION FOLLOWING NONDIFFERENTIAL REINFORCEMENT WITH DIFFERENTIAL AFFERENT CONSEQUENCES

By B. A. BILLINGSLEY, W. E. FEDDERSEN, and M. E. BITTERMAN,
University of Texas

In a recent experiment by Bitterman and Elam, rats were trained in the jumping apparatus to discriminate between horizontally and vertically striped cards after varying amounts of nondifferentially reinforced experience either with the two striped cards or with two identical gray cards.¹ Rate of learning was found to decrease with increasing amounts of nondifferentially reinforced experience of both kinds, although the relationship was significantly more marked in the case of the animals exposed to gray cards. Apparently a considerable degree of perceptual differentiation took place in the course of nondifferentially reinforced experience with the striped cards which offset in part what must be considered to be the general retarding effects of nondifferential reinforcement. In the present experiment the possibility of minimizing these retarding effects was explored. A situation was employed in which the difference between a pair of nondifferentially reinforced stimulus-cards was consistently linked with a striking difference in the afferent consequences of response (goal-box color) which had in itself no motivational significance. The hypothesis was tested that this linkage would maintain an orientation to the difference between the cards and thus facilitate the subsequent discrimination between them.

Subjects. The Ss were 42 experimentally naïve white rats, bred in the laboratory. They were 3-4 mo. of age at the beginning of the experiment.

Apparatus. A two-window jumping apparatus was employed. Behind the windows was a feeding platform divided by partitions into three goal-boxes, of which the first and third were painted black and the second white. The position of the platform relative to the windows could be so changed that it was possible to have either a white box behind the left window and a black box behind the right or a white box behind the right window and a black box behind the left. Below the windows was a cloth net into which the animals fell after responses to locked cards. With the exception of the goal boxes, the entire apparatus was mid-gray in color.

Four stimulus-cards were used, of which two were identical gray cards (slightly darker in color than the gray of the rest of the apparatus for the purpose of contrast with the surround) and two were black-and-white vertically striped cards differing in stripe-width. The narrower stripes measured $\frac{1}{4}$ in. and the wider stripes $\frac{1}{2}$ in. Since there were twice as many narrow white stripes as wide white stripes, the two cards were equal in total brightness. In previous experiments in this series,

* Accepted for publication May 20, 1953.

¹ M. E. Bitterman and C. B. Elam, Discrimination following varying amounts of nondifferential reinforcement, this JOURNAL, 67, 1954, 133-137.

the critical stimuli were striped cards differing in direction (horizontal-vertical) rather than width.² The more difficult width-discrimination was introduced in the present experiment in an attempt to weight more heavily the perceptual aspect of the final problem.

Preliminary training. The Ss were placed on a 24-hr. feeding schedule and maintained on that schedule throughout the experiment. After adjustment to handling they were fed wet mash equally often in white and black goal-boxes over a period of several days. Then they were trained to jump gradually increasing distances (to a maximum of 9 in.) from the jumping platform through the open windows, with each jump being rewarded by a nibble of wet mash. By shifting the positions of black and white goal-boxes and by a system of alternating free and manually guided trials, frequency of response to black and white as well as to right and left was equated. On the basis of preferences for position (which were quite marked) and for color (which were relatively slight) manifested during the preliminary training, the Ss were divided into three matched groups of 14 animals each.

Experimental design. During the first stage of the experiment, which lasted 12 days, all groups were given eight nondifferentially reinforced trials per day. Group I jumped to the pair of gray cards. The alternation of free trials with manually guided ones ensured that each S would jump equally often to the right and to the left, and appropriate shifts of the feeding platform ensured that S would be reinforced equally often in black and white goal-boxes. Group II jumped to the two striped cards presented equally often in each of the two simultaneous configurations (wide-narrow and narrow-wide). The system of free and guided trials provided equal experience with each card and each window. Half the responses to each card and to each window led to reinforcement in the black goal-box, while the remaining responses led to reinforcement in the white goal-box. The conditions for Group III were identical with those for Group II, except that there was a consistent relation between stripe-width and goal-box color. For half the Ss in Group III, response to wide stripes led always to reinforcement in the black box and response to narrow stripes led always to reinforcement in the white box; for the remaining Ss in the group the opposite relation between width and color obtained.

In the second stage of the experiment, all Ss were trained to discriminate between the two striped cards. Half the Ss in each group, selected at random, were reinforced for responding to the wide stripes, and the remaining Ss were reinforced for responding to the narrow stripes. Since the incorrect card was locked in place, each response to it was terminated by a fall into the net. Correct responses were for each S always rewarded in a box of the same color. Half of the Ss in Group I and Group II, selected at random, were reinforced in the white box, and the remaining Ss in each group were reinforced in the black. For each S of Group III (in which there had been a consistent relation between width and color in the first phase of the experiment) color of the goal-box was automatically determined by the selection of the positive stripe-width; the color was the one which had pre-

² Bitterman, A. D. Calvin, and C. B. Elam, Perceptual differentiation in the course of nondifferential reinforcement, *J. Comp. & Physiol. Psychol.*, 46, 1953, 393-397; Bitterman, Elam, and E. C. Wortz, Perceptual differentiation as a function of nondifferential reward and punishment, *J. Comp. & Physiol. Psychol.*, 46, 1953, 475-478; Bitterman and Elam, *op. cit.*, 133-137.

viously been associated with that width. All *Ss* were given eight trials per day by the method of correction. A maximum of three errors was permitted on each trial, following which *S* was guided in the correct direction; a correct jump terminated each trial. The criterion of learning was two consecutive errorless days.

Results. During the unguided trials of the first stage of the experiment, all *Ss* responded primarily in terms of preferences for position and no *S* gave evidence of a significant preference for stripe-width. In Table I, the performance of the three groups in the second stage of the experiment is summarized in terms of mean initial errors, mean total errors, and mean days to criterion. The differences among the groups were very small, and none approached statistical significance.

TABLE I

PERFORMANCE OF THE THREE GROUPS IN THE DISCRIMINATION OF STRIPE-WIDTH

Index	Group I	Group II	Group III
Initial errors	21.2	22.2	21.6
Total errors	37.4	37.1	39.4
Days to criterion	8.1	9.1	9.2

The close similarity of the performance of Group III to that of Group II contradicts the hypothesis which the experiment was designed to test; namely, that differential afferent consequences of response to the two nondifferentially reinforced stimuli would facilitate subsequent discrimination between them. It is conceivable, however, that this outcome is due merely to an unfortunate choice of stimuli, a conclusion suggested by the complete uniformity of behavior in all three groups. Supplementary tests demonstrate that the white rat fails after weeks of training to discriminate between stripe-widths on the order of those employed in this experiment even under conditions of *differential* primary (food) and secondary (goal-box color) reinforcement when no punishment for errors is administered—that is, when incorrect response results in admission to an empty goal-box rather than in an encounter with a locked card and a fall into the net. In view of these findings, it is perhaps not surprising that no perceptual differentiation at all seemed to take place in the first stage of the experiment. The hypothesis in question cannot, therefore, be rejected until a test is made with more appropriate stimuli. In general, work on the problem of non-differential reinforcement must be restricted to relatively dissimilar stimuli which are capable of discrimination on the basis of differential reinforcement alone.

Summary. For one group of animals trained in the jumping apparatus, the difference between a pair of nondifferentially reinforced stimulus-cards was consistently linked with a striking difference in the afferent consequences of response (goal-box color), a procedure designed to test the hypothesis that such linkage would maintain an orientation to the difference between the cards and thus facilitate the subsequent discrimination between them. Comparison of the performance of the group so treated with the performance of two control groups provided no support for the hypothesis.

RECOGNITION AND THE LOCALIZATION OF VISUAL TRACES

By HANS WALLACH and PAULINE AUSTIN, Swarthmore College

If an unknown figure is shown in one retinal position, will this figure be recognized later when it is shown in another retinal position? This question was answered in the affirmative more than fifty years ago by J. von Kries and the experiments of Becher confirmed Von Kries' position,¹ but psychologists have continued to seek some effect of retinal position on recognition. Stimulation of the left half of the retina activates the striate area of the left hemisphere and stimulation of the right side of the retina activates the right hemisphere. If a figure is first presented in such a way that its image falls on the left half of the retina and later is projected on the right half of the retina, the path of neural excitation to the cortical locus of the first process certainly is different from what it would be if the second presentation were made in the same retinal location as the first one. If memory traces are localized at the cortical place of the primary process which they represent, or if that place is in some other fashion involved in their arousal, such differences in the path of neural excitation should indeed exist. There is, of course, still the question whether such a difference, if it exists, can be detected. We know of four at-



FIG. 1. THE AMBIGUOUS FIGURE

tempts to show an effect of retinal position on recognition. They all failed and the results of the work therefore remained unpublished. Our own attempt which is here reported was made because we felt that we could employ an unusually sensitive method.

We used the silhouette of an ambiguous figure which was first used by Irvin Rock in as yet unpublished research on spatial orientation. This figure is usually recognized as a dog when it is presented with the long axis in horizontal orientation, and as a chef when that axis is vertical. Tilted 45° as in Fig. 1 it is a fairly balanced ambiguous figure. Our experiment consisted in presenting the dog-version in one retinal position and the chef-version in another one. Finally, the figure was

* Accepted for publication June 29, 1953.

¹ Wolfgang Köhler, *Dynamics in Psychology*, 1940, 128-129.

presented in its ambiguous orientation in either one of these positions. It would then presumably be recognized either as a dog or as a chef. An effect of position would be demonstrated if the ambiguous version were recognized predominately as that figure which had previously been given in the same retinal position.

In an experiment of this kind, the critical figure must be presented in such a way that only spontaneous recognition occurs, and no expectation on the part of *O* related to the purpose of the experiment can be allowed to develop. We therefore cast our problem in the form of an experiment on peripheral recognition of familiar forms. Nine silhouettes of familiar forms were presented singly in four different positions a short distance from a fixation mark. Among these were the dog-version and the chef-version of the critical figure. The tenth figure presented was the critical figure in its ambiguous orientation. With ten figures presented for recognition, there was little danger that *O* would concern himself specially with one of the critical figures.

Procedure. Each figure was flashed on a screen by an opaque projector for 0.4 sec. The size of the figures on the screen varied from 3.0 in. to 4.5 in. for the larger dimension. The fixation-mark consisted of a light spot of $\frac{1}{4}$ -in. diameter (projected on the screen by another lantern) which was 53 in. above the floor and remained visible throughout the experiment. Each figure was centered about a point 2.5 in. from the fixation-mark at an angle of 35° above or below the horizontal. It was always presented in one of the four positions, upper left, lower left, upper right or lower right, except for the critical figures whose position varied. The *O* was seated at a distance of 53 in. from the screen directly in front of the fixation-mark. The room was dimly illuminated by a table lamp placed at the far end behind *O*.

The following instructions were given: "This is an experiment on how people recognize things. I will flash some pictures on the screen very quickly. Looking directly at the point of light, will you name the object which you see? Remember, always look directly at the point of light during the exposures. I will give you a ready signal before each picture."

The non-critical figures used represented the following things: whale, tree, car, duck, house, pitcher, and boat. The critical figures were presented in the second and sixth places in the sequence—after the whale and after the duck—and in the tenth place. They were given only in the upper positions. When the chef-version was given on the left, the dog-version was given on the right, and *vice versa*. The presentation of the critical figures varied in three ways. The dog-version was given either on the left or on the right, and it was given either in the second or the sixth place in the sequence; the ambiguous version was given either on the left or on the right. We used all combinations of these variations in a balanced design which thus involved eight different presentation-sequences.

Altogether 87 *O*s, undergraduates of Swarthmore College, were used. To 39 of them, however, the ambiguous figure was not presented, either because they had not recognized one of the versions of the critical figure (18 *O*s), or they had incorrectly recognized it (4 *O*s), or they had seen it as the other version of the critical figure (17 *O*s). Apparently the dog-version was easier to recognize than the chef-version, because 37 of the 39 failures of recognition occurred when the latter was presented. When an *O* failed to recognize one of the critical figures, the presentation sequence employed was given to another *O*. This procedure was continued

until each of the eight presentation-sequences had been completed six times and 48 *O*s had completed the whole experiment.

Results. All 48 *O*s saw the ambiguous form of the critical figure either as a dog or as a chef, and as one of them only. Of the 24 *O*s to whom the ambiguous figure was presented in the retinal position in which the dog-version had previously been given, 16 recognized the ambiguous figure as a dog and 8 as a chef. Of the remaining 24 *O*s, to whom the ambiguous figure was presented in the position in which the chef-version had previously been given, 20 recognized the ambiguous figure as a chef and 4 as a dog. Thus, 36 of the 48 *O*s recognized the ambiguous figure as the version which had previously been presented in the same retinal position (*near-plant*), and 12 *O*s recognized the ambiguous figure as the version which had been presented on the other side of the fixation-mark (*far-plant*). The difference is significant beyond the 0.2-% level.

Before we draw the conclusion that retinal position of stimulation affects the recognition of a figure, an alternative interpretation must be discussed. When the 12 cases in which recognition had apparently been in terms of the *far-plant* were inspected, it was found that in 9 of them the ambiguous figure had been presented on the left of the fixation-mark. In the majority of these cases, then, the version in terms of which the ambiguous figure was recognized (the effective plant) had been presented on the right of the fixation-mark and thus had caused a primary cortical process in the left hemisphere. Although the difference between 9 and 3 does not represent a reliable trend, the possibility that a plant in the right half of the visual field proves more effective than a plant in the left half must be seriously considered.

Among the results for all 48 *O*s there were 30 cases in which what turned out to be the effective plant had been presented in the right half of the visual field. The question arises as to whether our main result can be explained on the assumption that a right-field plant dominates recognition. To answer this question we must consider the 18 cases in which, because the effective plant had been on the left side of the fixation-mark, the results could not possibly be ascribed to left hemispheric dominance. Of these 18 *O*s, 15 recognized the ambiguous figure in terms of the *near-plant* and this preponderance is statistically reliable at the 1-% level. Our results thus show an effect of position even if all cases of effective plant in the right half of the visual field are assumed to be the consequence of left hemispheric dominance. Whether such dominance does play a rôle in our experiment remains to be seen; our results do not as yet demonstrate it reliably.

THE EFFECT OF TWO BRIGHTNESS FACTORS UPON THE RATE OF FLUCTUATION OF REVERSIBLE PERSPECTIVES

By HELEN K. MULL, NANCY ORD, and NAN LOCKE, Sweet Briar College

In this study an attempt was made to discover the effect of two brightness factors upon the rate of fluctuation of reversible perspectives. The two factors considered were: (1) brightness contrast between contour and background; and (2) illumination.

The Necker cube was used as the stimulus-object. It is perhaps questionable whether this figure furnishes the best instance of what Köhler calls contrast between figure and environment, which, together with neighboring areas produce, as he believes, a current in tissue in the visual brain center, the strength of which depends on the brightness difference between them.¹ The cube was chosen, however, because it was thought to furnish changes defined more sharply than any other figure.

Observers and apparatus. Thirty-six students (women from Sweet Briar College), divided into six groups of six each, served as Os in this study. A 2-in. cube, drawn with dark gray contour lines $\frac{1}{8}$ in. wide, was presented under three conditions: (A) on a white ground, under 200-w. illumination; (B) on a medium gray ground, under 200-w. illumination; and (C) on a white ground, under 15-w. illumination. (A) and (B) furnish a difference in degree of contrast, with illumination constant; and (A) and (C) provide a difference in amount of illumination, with contrast constant, as both outline and ground are here reduced by the same fraction.

Köhler holds, as will be remembered, that the processes of the visual center tend to block themselves, thus producing a shift in perspective; and, further, that the rate of blocking depends upon the strength of the current which, in turn, depends upon the amount of brightness contrast between figure and environment.² If this theory is true, then the greater the brightness contrast, the faster should be the rate of fluctuation. It was this point that the present experiment was designed to test under conditions which measured the efficacy of contrast apart from degree of illumination, and of illumination apart from contrast.

Procedure. After preliminary instruction, O, with chin in head-rest and seated 3 ft. from the pattern, closed her eyes for 15 sec., then fixated a neutral point at the center of the cube and reported the shifts in perspective as they appeared until 10 changes occurred. The time of every change was noted. This procedure, repeated 20 times, constituted a series of observations under one of the conditions described above. After a 2-min. rest, a similar series was made under each of the other condi-

* Accepted for publication June 25, 1953.

¹ Wolfgang Köhler, *Dynamics in Psychology*, 1940, 73 ff.

² Köhler and Hans Wallach, Visual after-effects: An investigation of visual processes, *Proc. Amer. Phil. Soc.*, 88, 1944, 328.

tions. To equalize temporal effects, these conditions were presented in counter-balanced order to every group of *O*s in each of the six possible temporal sequences: viz. ABC, ACB, BAC, BCA, CAB, CBA. All observations were made in a darkroom, with the light 3 ft. above *O*'s head.

Results. In Condition (A) the average length of the twenty 10-shift periods for the 36 *O*s was 30.6 sec.; in (B), 29.5 sec.; and in (C), 30.0 sec. The *CR* between the averages of (A) and (B) is 0.31, with 20 of the 36 *O*s showing the general direction of difference. For (A) and (C), the *CR* is 1.65, and 21 *O*s yielded results in the same direction. Rates of shift are, practically without exception, highly irregular under all conditions, both for the individual *O* and among the different *O*s.

The most obvious statements that can be made regarding our results are: (1) neither a difference in amount of contrast nor of general illumination influences the rate of fluctuation; and (2) neither factor is more effective than the other.

In explanation of these results it is tempting to suggest that they furnish some evidence against the adequacy of a single factor, such as blocking in visual brain areas, to account for them because this factor requires a greater dependency of rate upon sensory conditions, as well as a greater uniformity and regularity of rate, than is here displayed. The results could be explained, however, by postulating a second factor, *i.e.* motor.³ This complicating factor might then be supposed to mask the effects of the processes of the visual center and thus account for the ineffectiveness of differences in stimulation, as well as for the irregularity and lack of uniformity of rates of fluctuation.

³ Cf. M. F. Washburn, H. Mallay, and A. Naylor, The influence of the size of an outline cube on the fluctuations of its perspective, this JOURNAL, 43, 1931, 484-489; also R. S. Woodworth, *Experimental Psychology*, 1938, 574.

APPARATUS

APPARATUS FOR RECORDING AUTONOMIC STATES AND CHANGES

By R. C. DAVIS, G. F. SIDDON, and G. L. STOUT, Indiana University

Three pieces of apparatus, devised to record autonomic states and changes, are described here. The first apparatus is an amplifier with an automatic reset for recording galvanic skin responses and similar changes that are slow; the second is an electrical impedance plethysmograph; and the third, a recorder for pulse and arterial pressure.

I. AN AMPLIFIER WITH AUTOMATIC RESET FOR GALVANIC SKIN RESPONSES AND SIMILAR CHANGES

Many of the autonomically controlled processes which respond to stimulation are relatively slow changing. The galvanic skin response is fairly typical. A large response will take several seconds to reach its maximum and perhaps four or five times as long to recede. With phenomena of this sort amplifiers of the usual type containing reactances are impractical because the reactances would need to be of such enormous size. For amplifying such changes one must resort to direct coupled amplifiers or to converting circuits in which the slow wave is made to modulate a carrier.

In direct coupled amplification an increase or decrease in the direct current component of the input is likely to force the indicator off scale. Ordinarily a manual adjustment is required to bring it back to a useful range. There is often loss of a critical point in the record while this change is being made. Furthermore, with multiple channel recording no operator could monitor all channels at once. An obvious solution to the problem is to use instruments of such low sensitivity that they seldom reach the limits of their scales, but this solution entails a loss of precision and ease of reading. An alternative is to introduce an automatic feedback device which will shift the zero point of the scale by a known amount whenever occasion requires, so as to make the device operate over a different range. The circuit for such an instrument is shown in Fig. 1.

* The apparatus described here were developed under Contract N6onr-18011 between Indiana University and the Office of Naval Research, U.S. Navy.

As shown in Fig. 1, the input signal is impressed between the ground and the grid of the first amplifier (sometimes through resistors of the sensitivity control). The connection must be so made that the negative side of the DC input is on the ground terminal. Then the signal and bucking voltage are in opposite directions in the grid-cathode circuit of the 12SH7, and the net grid-cathode voltage is the difference of the two, plus whatever is provided by the two 3 K resistors shown in the cathode lead. When there is no input and no bucking voltage one of these (known as the 'zero set') is so adjusted that the 12SH7 has its proper operating bias (with grid negative to cathode). When a signal is introduced the bucking voltage will be adjusted (by the device described below) until the tube has that amount of bias, or something near enough to it to be in its proper operating range. If a wave, positive or negative, appears in the signal the tube will transmit this.

The series of 500 K resistors across the input serves as a sensitivity selector. These are mounted on a rotary switch which is manually operated. The sum of the 10 resistors is so large that no appreciable current is drawn by the device. According to the position of the switch a certain number of tenths of the whole signal is put into the amplifier. Of course the value of the whole can be obtained by appropriate multiplication.

The amplifier has two second stages, each composed of a 6SN7. The first of these serves to amplify the signal to provide the output of the system. The second amplifies the signal to provide the feedback required for the automatic reset of the bucking voltage.

In the first of these stages the two halves of the 6SN7 are connected in quasi push-pull fashion. If only one half were used it would be difficult to arrange for it to operate on a linear part of its characteristic. The other half therefore is used to oppose the one and improve the linearity. The output meter (0-1mA) and the external output connections are located between the two plates of this 6SN7. Only one of the grids, however, is connected to the signal, the other being grounded.

In the other 6SN7 the output of each half is used to operate a relay which in turn causes the reset motor to run forward in the one case and backward in the other. The signal from the 12SH7 is given to both halves of this tube. When the instrument is originally set up the cathode bias on the first half is adjusted so that the tube conducts and throws its relay ('low' relay) whenever the output meter comes on scale from below. Since the relay is connected to break a circuit this means that the motor will run, in the appropriate direction, until there is an output reading on the meter; then it will drop out. The other half of the 6SN7 has a higher cathode bias and will not conduct enough to move its relay until the output meter reads full scale, *i.e.* 1m.amp. It then throws the 'high' relay which has a 'make' connection and causes the motor to run in the other direction. Since the motor turns the rotary switch which regulates voltage on the input grid this mechanism acts as step-wise negative feedback.

Two specially troublesome effects were encountered when these devices were first tried. With certain 12SH7 tubes there was a tendency for the grid to draw current even when it was negative, and when there is grid current the amplifier becomes non-linear. The remedy turned out to be proper adjustment of the screen voltage. A variable resistor for this purpose is shown in the diagram.

The rotary switch used in the reset mechanism must be of the 'shorting' type;

it must have its wiper in contact with some terminal at all times because otherwise the system will oscillate. Poor quality switches originally gave a good deal of trouble by frequently leaving an open contact.

For experiments in which only one variable is to be recorded (one channel of GSR) the output of one of these amplifiers may be put into an Esterline-Angus recording milliammeter or similar instrument, 0.1 m.amp. sensitivity. For experiments requiring multiple recording, further amplification may be needed to drive the power amplifiers which move the stylus. It is also desirable that the signal should have a ground connection—which it does not. For this purpose a converter amplifier can be introduced following the direct coupled amplifiers. The circuit for this is shown in Fig. 2. The signal is converted to a modulated 60~ wave by means of a vibrator (without a ground connection).¹ After two stages of amplification the

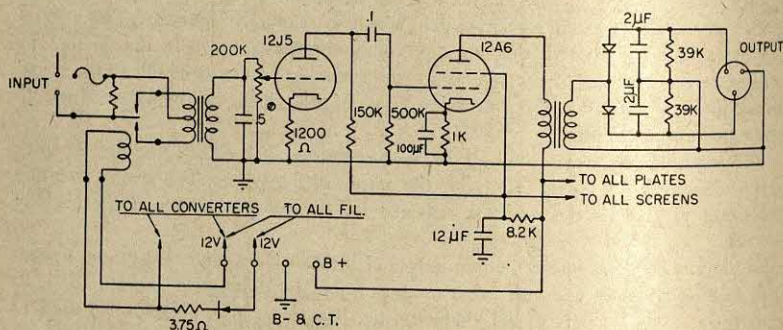


FIG. 2. CONVERTER AMPLIFIER TO CONNECT OUTPUT OF DIRECT COUPLED AMPLIFIER TO CENTER GROUNDED RECORDING DEVICE

signal is rectified and put into push-pull form with a center ground. This arrangement gives the type of input desired for the power amplifiers.

To use this self adjusting amplifier for recording the GSR (resistance change method), it is necessary only to supply a regulated current through *S* and connect him directly to the input of the amplifier. Fig. 3 is a schematic diagram of a circuit supplying the electrotonic current through *S*.² Voltage for supplying the current is derived from the 6.3-v. winding of the transformer in a direct coupled amplifier (see p. 343). This is stepped up by means of a small filament transformer used in reverse. The current is then rectified and applied to the circuit consisting of *S* and the tube 6SH7, whose function is to keep the current constant. (*S* is in the plate circuit of this tube.)³ By means of a double throw switch, *S* may be replaced by a

¹ Ordinary vibrators adapted to these circuit requirements were found to get out of adjustment very easily. A 'Brown' converter proved satisfactory.

² Basically the circuit is that described by R. C. Davis, and J. M. Porter, Jr., A measuring device for the galvanic skin reflex, *J. Gen. Psychol.*, 5, 1931, 115-120.

³ The electrodes used with this apparatus consist of plastic cups about 1 in. in diameter. In the bottom of each is a plate of zinc to which the lead wire is soldered. Above this is a layer of manganese dioxide (MnO_2) soaked with normal salt solution, and on top of that in contact with the skin is a felt pad soaked in normal salt. *S*'s skin is rubbed with electrode jelly prior to attachment of electrodes.

'stand-in' for purposes of calibration. This stand-in, as seen in the diagram, consists of a series of ten 500- Ω resistors, any number of which may be put into the circuit by means of a rotary switch.

In this circuit, since the current is constant, the voltage across S is proportional to his apparent resistance. Consequently a measurement of this voltage is a measurement of S 's resistance. The voltage therefore is the signal which is carried to the input of a direct coupled amplifier with an automatic reset device (see p. 343).

In operation the amplifier will automatically adjust itself to its proper operating

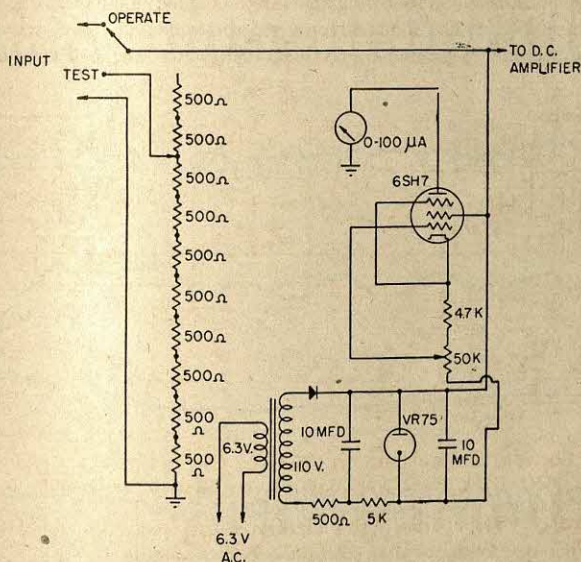


FIG. 3. CIRCUIT FOR SUPPLYING REGULATED CURRENT TO S FOR GSR MEASUREMENT WITH DIRECT COUPLED AMPLIFIER

range as soon as it is connected to *S*. When it settles on this range, with the indicating pointer on the dial, the index number showing through a hole in the panel will indicate how many steps the rotary switch has advanced from zero. If the value of each step is known, the index number therefore shows the total voltage which the mechanism has applied to offset the *S*'s apparent resistance.

Once S has been inserted into the circuit and the mechanism has produced a reading on scale, the automatic reset will keep it there. When the indicator reaches either end of the scale the mechanism will throw it back on scale again by adding or subtracting a constant voltage value. Thereupon the index on the panel will show a new number, indicating a new base level of resistance. On a graphic record taken from the instrument such a quick jump in level is easily distinguished from the slow changes in GSR (see upper line in Fig. 4). Hence, if one knows the base-level at the beginning of a run, he can ascertain it at any point in the record by allowing for the appropriate number of resets in each direction.

II. AN ELECTRICAL PLETHYSMOGRAPH

The plethysmograph has long been used in psychology. The inconvenience and inaccuracy of the hydraulic and pneumatic models has, however, led to its neglect. This is unfortunate since the constriction-dilation of the

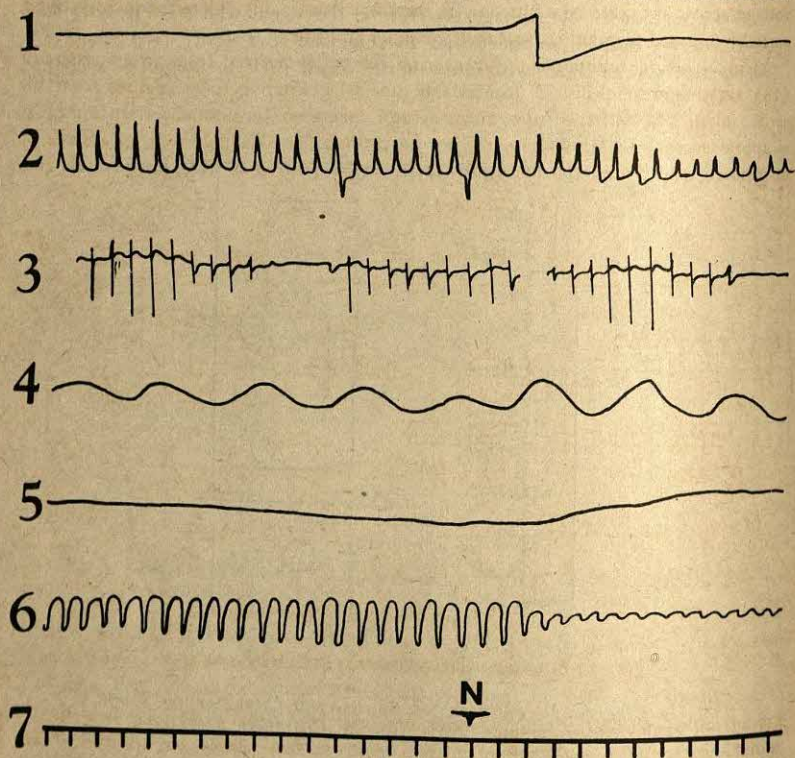


FIG. 4. SECTION OF RECORD PRODUCED BY APPARATUS DESCRIBED

1, GSR showing operation of automatic reset when GSR reaches limit of scale; 2, pressure pulse; 3, pressure pulse with periodic occlusion by pressure cuff operated by mercury pump; 4, chest breathing recorded from movement of variable arm of small potentiometer strapped around chest; 5, DC recording of finger volume. Upward movement represents decrease; 6, finger volume pulse, and 7, time in seconds (N indicates noise stimulus).

small vessels is perhaps the most important variation in the circulatory system.

Nyboer has described and validated a plethysmograph which depends upon electrical impedance.⁴ His device, however, is needlessly elaborate and frequently it

⁴J. Nyboer, Plethysmograph: Impedance, in O. Glasser's, *Medical Physics*, 2, 1950, 736-743.

cannot be used because the electrical properties of the skin were not those assumed in the design.

We have, therefore, revised his circuit. The basic plan is to apply a high frequency current across the part whose volume is to be measured. The impedance to the current should vary with the volume, and the circuit is arranged so that the RF voltage across the part is a function of the impedance. This voltage is then rectified, partially opposed by another, and the remainder amplified.

As in Nyboer's apparatus, the measuring current is derived from an RF generator (135 kc.) shown in Fig. 5 around the tube 6J5. From here we depart from the earlier plan. The voltage from the generator taken off through the transformer T_1 is applied across S and a series resistor of 1000 Ω . This voltage will, of course, be distributed over S and the fixed resistor, R , in proportion to their impedances. If S

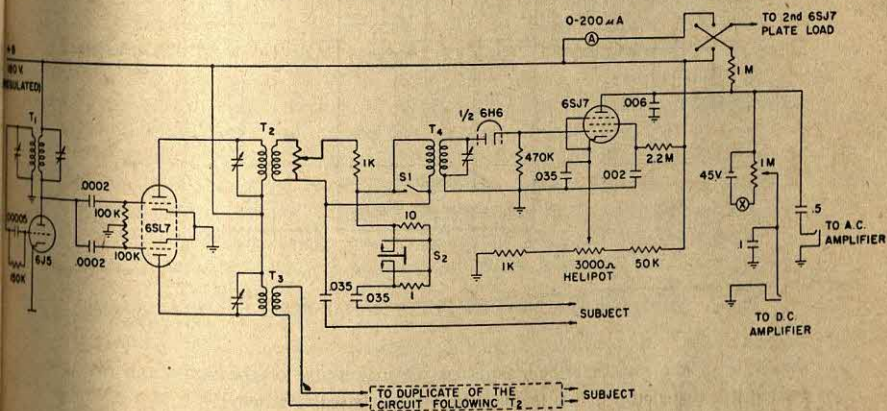


FIG. 5. ELECTRICAL IMPEDANCE PLETHYSMOGRAPH

has an impedance Z the voltage across him will be $[Z/(Z + R)] \cdot E$ — E being the applied voltage. Z being small in relation to R , we may simply say that the voltage across S is proportional to Z , his impedance. To eliminate the effects of phase-shift, the next operation of the circuit is to rectify this voltage by the tube 6H6. Then, because the voltage is quite large with respect to its fluctuations, in which we are interested, a part of it is opposed by a voltage taken from the cathode circuit of the next tube (6SJ7). The algebraic sum of these voltages is then put into the 6SJ7 for amplification. The meter located in the output of this tube indicates how the opposing voltage in the input circuit should be adjusted. The size of the output current which corresponds to its optimal input bias was determined when the instruments were constructed. For one of the instruments, for example, it turned out to be 160 μ /amp., hence, when an S was first introduced into the circuit, the operator adjusted the opposing voltage until the meter read 160 μ /amp. When readings of total impedance were desired this was done by manipulation of a calibrated potentiometer (Helipot) shown in the diagram. This setting will not need to be changed during a sitting in most cases.

The switching arrangement marked S_2 provides a convenient calibration. By throwing the switch in one direction or the other a 1 or 10 Ω resistor is removed from the circuit.

The potentiometer P_1 is provided to reduce the measuring voltage applied to S . If the voltage drop across S is large, it will require more neutralization from the Helipot than it can provide without interfering with the operation of the amplifier 6SJ7. The range of the instrument would be increased if a voltage higher than the indicated 180 were applied to this tube.

At the bottom of the diagram is shown a connection to a circuit which is the

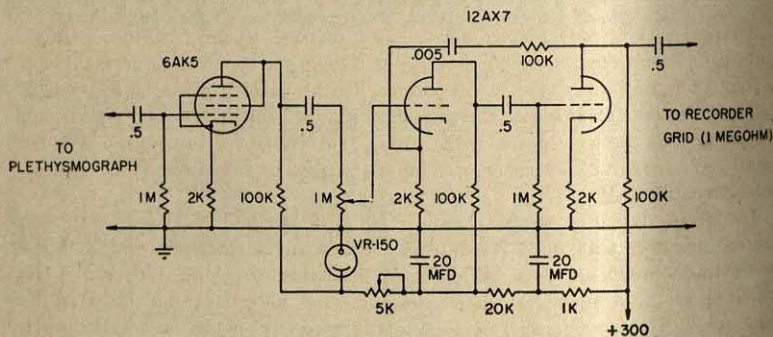


FIG. 6. AMPLIFIER FOR VOLUME PULSE SIGNAL FROM PLETHYSMOGRAPH

duplicate of the one described. This circuit makes use of the same high frequency source to record impedance from some other part of the body.

From the output of the 6SJ7 there are divergent paths which can be operated at the same time. One is a direct coupled pathway which will record all changes occurring in S 's impedance. (It is most useful for long term drifts, and might be more convenient if the pulse variations were eliminated by a shunting condenser.) Since these slow changes require further amplification they are passed on to a direct coupled amplifier such as described above. Before the output of the 6SJ7 could be handled by these, however, another neutralization had to occur to eliminate a large unvarying portion of the output. This is the purpose of the bucking battery working into a potentiometer shown in the diagram. This potentiometer setting constitutes a permanent adjustment so long as the battery and tube remain the same.

On occasion, however, we are interested in the pulse-wave itself, rather than in the slow impedance-drifts. For this purpose it is more convenient to use capacity coupling than direct coupling in subsequent amplification. Hence the second pathway out of the plethysmograph. Through a coupling condenser it leads to a voltage amplifier of standard design as shown in Fig. 6. The time constant of this amplifier was a matter of compromise. For the purposes in hand the circuit-values shown seemed best since they give a reasonably accurate pulse-wave and not too slow a recovery time. A record of the pulse component is shown just above the time line in Fig. 4, and of the DC component in the line above that.

Correlations obtained by H. A. Dawson, Jr., between such records and records from a hydraulic device give reason to believe that the two types differ only by their errors of measurement.

III. PULSE AND ARTERIAL PRESSURE

The pressure variation in the arterial pulse has properties quite different from those of the volume pulse of the finger, and it therefore constitutes an additional autonomic variable. Since the pulse in an artery is little affected by local conditions, it provides opportunity for pressure measurements.

A crystal pick-up was adapted for use as a pressure sensitive transducer. In the particular pick-up used, a crystal about $\frac{1}{2}$ in. square was mounted horizontally with a point projecting downward at one corner. To this point a plastic disk 5 mm. in diameter was cemented. In use the pick-up is strapped snugly around S's wrist with the plastic disk resting on the radial artery. The crystal, of course, resists movement, and such little movement as it allows should be indicative of the pressure variation on the disk.

The electrical output of the crystal is led to a standard resistance—a capacity coupled amplifier with good response in the low audio range and with sufficient sensitivity to produce a 1-cm. deflection of the stylus for a millivolt input. (This amplifier may also be used as an electrocardiograph since the requirements are the same for such an instrument.) A gain control makes it possible to adjust the size of the output. The second line from the top in Fig. 4 is a record of the pressure pulse taken with this equipment.

In measuring arterial systolic, diastolic, and mean pressures, it is necessary to have a means of occluding an artery with a known pressure and a pulse detector which will show when the artery is undergoing occlusion. The pressure pulse pick-up may be used for the latter purpose. When so used it plays a rôle like that of the stethoscope in the usual clinical measurement.

It is possible by means of a feed-back to apply a continuously occluding pressure to the upper arm—a pressure which varies directly with the physiological variations in blood pressure. Such a device prevents simultaneous reading of diastolic and systolic pressures, and is probably undesirable to use over long periods, since it necessarily blocks completely the circulation in the arm. Sometimes, therefore, intermittent pressure may be preferred. Hence a device was constructed for this purpose.

The general plan is so to apply pressure to a cuff that the pressure on the artery would increase gradually in a linear fashion to some point beyond the occlusion of the artery, to have the pressure then released and, after a brief pause, applied again. This cycle is repeated every 20 sec. Somewhere during the ascent of the pressure the pulse below the cuff would be cut off. Knowing when this happened from the continuous record, it is possible to read the corresponding point on the linearly ascending curve of pressure which is the point of systolic pressure at the moment.

A simple pump to accomplish this was constructed as follows. A reservoir of mercury, connected by tubing to the pressure cuff, was so mounted on a lever arm, which was rocked by a motor-driven cam, that the mercury flowed back and forth

between the reservoir and cuff.⁶ As the mercury flowed into the cuff it produced the desired pressure and as it flowed out it gave the desired release. The cam of the pump also operated two switches which indicated the beginning and the end of the pressure cycles.

A typical record is shown in Fig. 4. In reading such records the diminishing peaks of pulse waves may be connected with a straight line, and the point of its intersection with the base line taken to represent the systolic pressure. The distance of this point from the beginning of the pressure application should be proportional to the applied pressure at the moment. There is reason to believe that diastolic pressure can also be read. As the pressure climbs the pulse will, after a point, begin to increase. With still greater pressure it decreases and disappears. The point of maximum pulse should represent the point at which diastolic pressure is just balanced out.

Recent results from our laboratory, however, show that, while the pressure pulse is a very useful measure, any currently known method for measuring arterial pressure creates very serious disturbances in a psychological experiment.

⁶ Since it is desirable to keep the volume of the cuff as small as possible, it was made of a 6-in. length of surgical drainage tubing, 0.5 in. in diameter. As far as possible, this tube was kept from expanding in an outward direction from the arm by an inelastic bandage. When the cuff was applied it was drained of mercury and completely collapsed.

AN INEXPENSIVE SQUARE WAVE STIMULATOR

By J. M. HARRISON and C. E. MAYER, Boston University

In a conditioning experiment now in progress direct electrical stimulation of the nervous system is used as the CS. This work requires the use of a stimulator which produces electrical impulses of adjustable and known characteristics. While a variety of electrical stimulators are commercially available they are often unsuitable and their price may be prohibitive for the small laboratory. The stimulator described in this paper has a number of advantages. First, it is relatively cheap, the total cost of components being about \$65. Second, it supplies either single square waves of variable duration or a series of such waves at a repetition-rate determined by an external oscillator. (The single square wave is particularly useful in psychological experiments involving the use of electrical stimulation.) Finally,

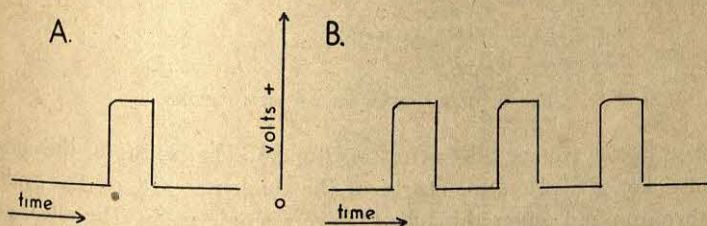


FIG. 1. OSCILLOGRAPHS OF THE OUTPUT OF THE STIMULATOR

A. Single square wave produced by pressing the key S1. B. Continuous output at 60 c.p.s. when 'time' control set to 6.5 msec.

the circuit is simple, requires few components and no critical adjustments. Oscillographs of the wave form are shown in Fig. 1.

Block diagram. A block diagram of the stimulator is shown in Fig. 2. The basic functional elements consist of the thyatron generator and its wave shaping (squaring) circuit. The thyatron generates a single pulse when a key is depressed. A series of pulses is produced when a sine wave is fed to the input of the pulse generator. The output of the stimulator is positive-going and has a maximum value of 12 v. on open circuit. The use of a cathode-follower in the output stage provides much better regulation of the output than can be obtained with a conventional output stage. Down to an external load of 2000 ohms the output voltage is virtually

independent of load. The pulse width and voltage of the output can be independently varied. Pulse widths from 0.5 m.sec. to 7.5 m.sec. can be obtained.

Circuit operation. Fig. 3 is a schematic diagram of the stimulator. The thyatron tube, $T1$, the two resistors, $R2$ and $R3$, and the condenser, $C3$, form the heart of the circuit. This is essentially a saw-tooth generator except that the thyatron is biased to prevent self-oscillation. This bias is

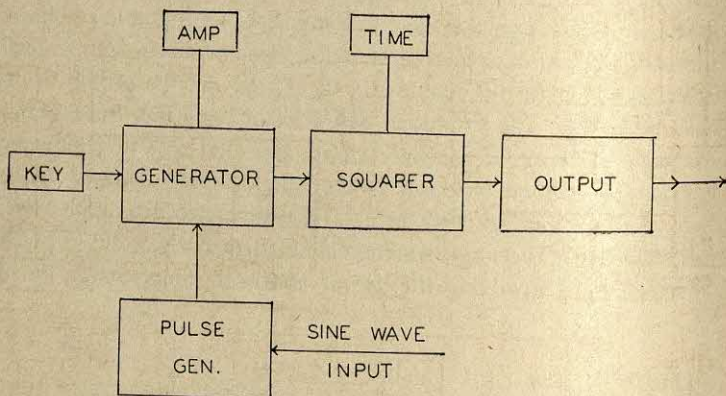


FIG. 2. BLOCK DIAGRAM OF STIMULATOR

supplied by the voltage drop across resistor $R5$. The closing of the 'single impulse' key, $S1$, places the charge of the condenser, $C1$, on the grid of the thyatron and causes this tube to give a single pulse. The alternative method of firing the thyatron is considered below.

The squarer circuit consists of a triode, $T2$, operating on an anode supply of about 30 v. and the resistors $R7$, $R8$, $R9$, and $R10$. The operation of the squarer depends upon the change in anode voltage of $T1$ exceeding the change in voltage required to swing the current of the tube $T2$ from its maximum value to zero. When the thyatron is not conducting, its anode is at the same voltage as the B -supply, 250 v. The anode of $T1$ is directly connected to the grid of the squarer, $T2$, via the series resistor $R7$. Thus $T2$ is taking its maximum current (saturation) when $T1$ is not conducting. When $T1$ conducts, its anode voltage will drop to about 30 v. This is a drop of some 220 v. from the non-conducting condition. Such a voltage excursion is much in excess of that required to reduce to zero the anode current of the squaring triode. After it has been discharged the timing condenser, $C3$, of the thyatron charges at a rate determined by the setting

of the 'time' control, $R3$. During most of the time of recovery of the anode voltage of the thyatron, the grid voltage of the squaring tube is zero. Finally, when the anode voltage of $T1$ approaches its final value, the anode current of the squaring tube rises to its maximum. This current change takes place rapidly because of the small voltage excursion of the grid of

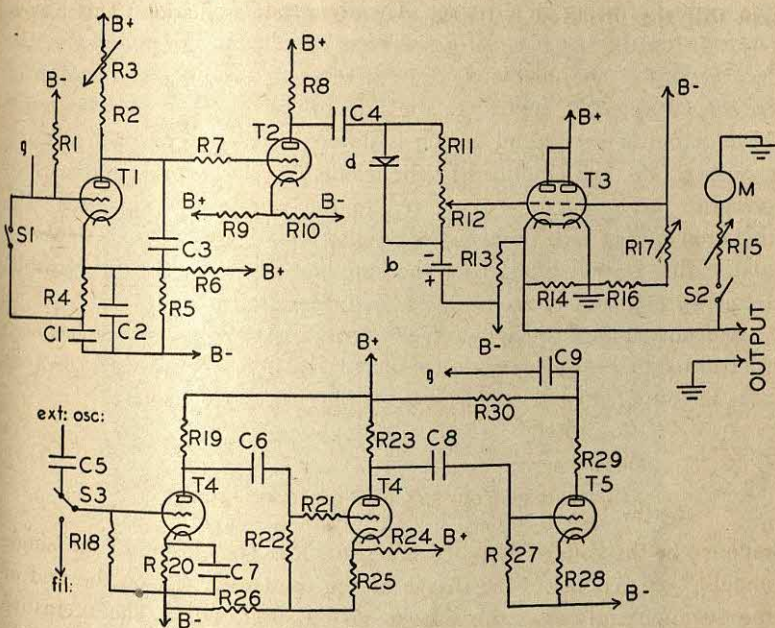


FIG. 3. WIRING DIAGRAM OF SQUARE WAVE STIMULATOR

$R1$, 100 K; $R2$, 11 K; $R3$, 250 K pot; $R4$, 1.5 M; $R5$, 4.7 K; $R6$, 20 K; $R7$, 0.5 M; $R8$, 0.25 M; $R9$, 3.3 K; $R10$, 22 K; $R11$, 10 K; $R12$, 0.5 M pot; $R13$, 6.8 K; $R14$, 4.7 K; $R15$, 100 K pot; $R16$, 195 ohms; $R17$, 200 ohms pot; $R18$, 0.5 M; $R19$, 100 K; $R20$, 3.9 K; $R21$, 100 K; $R22$, 0.5 M; $R23$, 100 K; $R24$, 3.3 K; $R25$, 220 ohms; $R26$, 22 K; $R27$, 200 K; $R28$, 2.7 K; $R29$, 40 K; $R30$, 220 K; $C1$, 0.002 mfd.; $C2$, 0.5 mfd.; $C3$, 0.015 mfd.; $C4$, 0.6 mfd.; $C5$, 0.015 mfd.; $C6$, 0.015 mfd.; $C7$, 100 mfd. electrolytic; $C8$, 0.0004 mfd.; $C9$, 0.02 mfd.; $S1$, $S2$, microswitches; $S3$, SPDT toggle switch; $T1$, 884; $T2$, 6L5; $T3$, $T4$, 6SN7; $T5$, 6C4; b , 7.5 v. C-battery; d , crystal diode 1N34; M , balance meter, 0-20 μ a.

$T2$ required for its anode current to go from zero to its maximum value. Thus, the output of $T2$ is a positive-going square wave whose duration is determined by the setting of the 'time' control. This square wave is fed via the 'amplitude' control, $R12$, to the grid of the output cathode-follower, $T3$. The second section of the output tube is arranged to have a cathode voltage equal to that of the first section when no signal is

being transmitted. This voltage is finely adjusted by means of the 'balance' control, *R17*, to produce a null reading on the meter, *M*. The output may be unbalanced in either direction to supply a polarizing current should it be required. After the stimulator has been allowed to warm up for 30 min., it holds its balance within 0.5 μ amp. over long periods.

The tubes *T4* and *T5* with their associated components form an alternative method of firing the thyatron. The thyatron discharges at a repetition-rate equal to the frequency of a sine wave supplied to the grid of *T4*. The two sections of the double triode, *T4*, serve as amplifier and squarer of the sine wave input applied to the grid of the first section. The square wave output of the second section is differentiated by the small coupling condenser, *C8*, and amplified by the triode *T5*. The narrow pulses which comprise the output of *T5* are fed to the grid of the thyatron. The thyatron fires at each positive-going pulse from *T5*. The input sine wave to the first section of *T4* may be supplied from an external variable-frequency oscillator or from the 60 cycle filament supply.

A neon stabilized *B*-supply of 250 v. is used. The stimulator has been constructed on a standard radio chassis 14 $\times$ 10 $\times$ 3 in. A 6 in. cover fits over the top of the chassis and protects the tubes from damage.

APPARATUS NOTES

ELECTRODES FOR MEASURING THE GSR IN SMALL ANIMALS

Various experimental situations exist in which it may be useful to obtain measures of autonomic activity in small animals, such as the laboratory rat. Such measures may be used, for example, in the study of 'emotional reactivity,' in equating the 'subjective magnitude' of stimuli from different sensory modalities, or in studies of conditioning and generalization.

The principal problem encountered in attempts to measure the GSR in small animals has been in the construction of electrodes which remain in constant con-

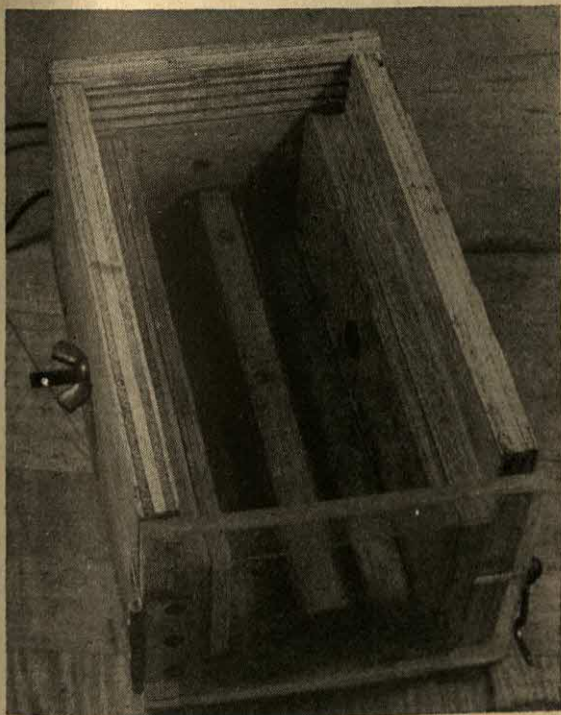


FIG. 1. AN ELECTRODE BOX FOR OBTAINING RECORDS OF THE GSR
IN SMALL ANIMALS

tact with appropriate areas of the skin, such as the foot-pads, yet do not interfere too much with the normal condition of S. Once satisfactory electrodes have been obtained, any measuring apparatus may be used which passes a slight current or

voltage through the *S*, which is treated as a variable resistor. The electrode-system described here has been employed successfully with laboratory rats, providing GSR-patterns similar to those obtained from human *Ss*. The procedure is to place the animal in a small box in which it stands on two strips of electrode-material (e.g., zinc or silver).

The box shown in Fig. 1 is 4 in. high, 5 in. wide, and 9 in. long. To keep *S* from lifting its paws from the electrodes, its movements in the box are restricted. The inside-height is adjustable by reason of the grooves in the inside walls which make it possible to insert the plastic top at varying distances from the floor; the inside-width and length of the box can be varied by inserting plywood panels along the walls and at one end of the box, holding them in place with a wing bolt. (For conditioning studies with a visual CS, it may be desirable to construct the walls, as well as the top, of clear plastic.) As an additional safeguard, a partition $\frac{1}{2}$ in. wide and 1 in. high is placed between the electrodes. This partition, which runs the length of the box, is cemented in place to serve also as insulation between the two electrodes.

In actual use, this method of obtaining autonomic measures is reasonably satisfactory. In a series of tests it was found that, after *S* became accustomed to the box, relatively stable records both of basal levels and of minor fluctuations in resistance could be obtained. In further tests, two *Ss* were adapted to a neutral stimulus (light), as indicated by the disappearance of the GSRs. The light was then paired with shock, and conditioned responses to the light were recorded.

University of Chicago
McGill University

ERNEST A. HAGGARD
WILLIAM R. THOMPSON

AN IMPROVED CIRCUIT FOR CONNECTING THE HUNTER TIMER AND SIMILAR TIMERS IN SEQUENCE

Hunter and Brown have described a circuit for connecting two or more of their electronic timers for automatic recycling operation.¹ This circuit utilizes an auxiliary relay, which serves mainly to avoid having the timers stop and immediately restart after reaching the end of a cycle. Such operation could introduce errors in timing because immediate restarting would not allow the timing-condenser in the RC circuit to acquire its full charge. The circuit shown in Fig. 1 eliminates this source of error without an auxiliary relay when three or more timers are used. Each timer in this circuit remains on until the next timer in the series shuts it off. Thus, when the last timer in a series reaches the end of its cycle, the first timer is ready to start from rest.

The sequence is begun by momentary closure of the starting switch (not shown) on any timer. Once started, a given timer, say the first, maintains itself through the NO side of its own 'start'-relay and the NC side of the 'finish'-relay on the next

¹ T. A. Hunter and J. S. Brown, A decade-type electronic interval-timer, this JOURNAL, 62, 1949, 570-575.

timer. The second timer is started through a pair of *NO* contacts on the finish-relay of the first timer. When the second timer reaches the end of its cycle, it shuts off the first one by breaking the circuit to the starting contacts of this timer through

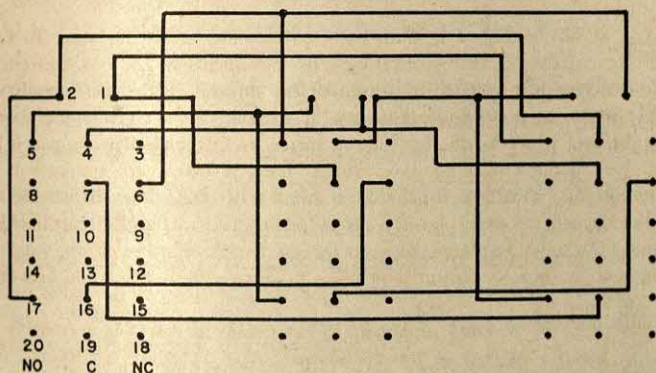


FIG. 1. CIRCUIT FOR OPERATION OF THREE TIMERS

Starting contacts, 1-2; finish-relay contacts, 3-12; start-relay contacts, 15-20.

a pair of *NC* contacts on the finish-relay of the second timer. The operation is similar for *N* timers. There are two sets of contacts on the finish-relay and one set on the start-relay of each timer available for external connections.

New York University

LAWRENCE KARLIN

NOTES AND DISCUSSIONS

MUSICAL MEMORY-MECHANISMS

In his interesting neurophysiological paper on "Memory Mechanisms," Penfield made several statements which may be criticized and elaborated psychologically for a better understanding of memory and consciousness.¹

Statement I: "We conclude that a man who has a good repertoire of remembered music, after he has selected one song, may hear it played through his own mind by causing impulses to pass down ganglionated nerve drains to the temporal cortex. The song will continue until he turns his attention elsewhere and shuts off those nerve impulses."

Comments and observations. (a) This memory-hearing occurs not only voluntarily, but also spontaneously, particularly after musical practicing or listening, and it may be so persistent that it can be suppressed only with difficulty. The voluntary memory-hearing is appreciated and even practiced by musicians. It can serve as a substitute for actual practicing and as a proof of perfect memorizing of a music piece for a concert performance. The persistent, repetitive memory-hearing (a kind of 'recurrent image') can be very disturbing to all kinds of activity, since it interferes with attention and requires considerable effort for suppression. (b) Even in this suppressed state the memory-mechanism may continue unconsciously (i.e. purely physiologically), and conscious memory-hearing may recur at progressive parts of the musical piece at times which correspond to the uninterrupted memory-hearing. This interrupted memory-hearing was accidentally observed by the author, when, after playing a sonata, he heard consecutively in memory parts of the first, second, and final movements.

The phenomenon was quantitatively studied on ten professional musicians by two different methods: (1) The O started memory-hearing of a familiar piece at the stop-watch time zero. The conscious memory-hearing was interrupted by conversation, reading, or looking at pictures. If conscious memory-hearing recurred, the watch was stopped. The piece was then played or heard by memory up to the same point and the time taken again. In eight musicians almost identical times were, as a rule, observed for interrupted memory-hearing and continuous hearing or playing. Occa-

¹ Wilder Penfield, Memory mechanisms, *Arch. Neurol. Psychol.*, 67, 1952, 178-198.

sional differences could be explained through omissions or duplications of a short phrase through mistakes at a point of deviation; but these errors occurred rarely. In two *O*s the music usually continued, where it became lost, if it continued at all. These *O*s definitely represented the exception to the rule. (2) A practical alternative to this method was the stop-watch timing of the complete piece, once continued, once or several times interrupted, but then voluntarily continued to the end. The results were the same, even for pieces with several movements and many points of deviation.

Unconscious (functional) activity of the memory mechanism also is indicated through the following experience of several *O*s: The memory-hearing may appear suddenly at any arbitrary point of a musical piece, indicating preceding unconscious activity of the memory-mechanism. No musician would start playing in the middle of a phrase; and most musicians could not do it if they wished.

From the observed facts one may conclude that the memory-mechanism is capable of reproductive activities, isolated from the center of consciousness in the centrencephalon (as postulated by Penfield and supported through the studies of Magoun and of Jasper²). A similar disconnection occurs in inattentive reading, when the motor activity of reading goes on while attention is paid to something else and intelligent understanding of the text is lacking until integrating connection with the centrencephalon is re-established. It appears that the memory-pathways, established by facilitation or *Bahnung* (Verworn), exist in two different states, *i.e.* dormant, and active engrams. Even the active state remains unconscious unless functional connection with the centrencephalon is established.

Statement II: "Whenever a normal person is paying conscious attention to something, he is simultaneously recording it in the temporal cortex of each hemisphere."

Comments and criticism: (a) The present writer is dubious about the exclusive, or even predominant rôle of the temporal cortex for memory-recording. Direct stimulation of the visuo-psychic area leads to the vision of moving objects, figures, or persons, *e.g.* menacing flames, fluttering butterflies, or "robbers with guns."³ These experiences must be considered as

² H. W. Magoun, An ascending reticular activating system in the brain stem, *Arch. Neurol. Psychol.*, 67, 1952, 145-154; Herbert Jasper, Cosimo Ajmone-Marsan, and Julius Stoll, Corticofugal projections to the brain stem, *Arch. Neurol. Psychol.*, 67, 1952, 155-171.

³ Penfield, *op. cit.*, 178-198; O. Foerster und Penfield, Der Narbenzug am und im Gehirn bei traumatischer Epilepsie in seiner Bedeutung für das Zustandekommen

recollective hallucinations with the same right as the ones elicited through direct stimulation of the temporal lobe. (b) Most of Penfield's results of temporal lobe stimulation are acoustical in nature. The few of a visual type could have been triggered through their association with acoustical components, directly obtained through stimulation of the temporal lobe. (c) While spatial recordings and integrations appear psychophysically related to the pulvino-parietal system (evidenced through asteriognosis in parietal lesions), and while acoustic and time-recordings and integrations appear similarly connected with the temporal lobe (evidenced by Penfield's results), these relations should not be considered as exclusive; overlapping functions should be postulated on general cerebro-physiological principles. A case in which musical memory and reproduction persisted after bilateral removal of the anterior and superior parts of the temporal lobes (quoted by Jasper during a discussion) supports this view. (d) The writer is dubious also about the necessity of conscious attention for memory-recording. For the great musical masterworks, attention to and even careful comparative study of their architecture is almost indispensable. Even here, however, many sensory and most motor-activities remain outside attention and below the level of consciousness. These facts represent musical examples of the difference between habit-formation and intelligent learning. If the 'memory areas' are responsible for habit, we must postulate recording without conscious attention. (e) Musical memory-mechanisms exhibit considerable variation according to the kind of acquisition. A simple piece of music, which can be heard by memory, and reproduced through humming, whistling or playing on any instrument and in any key ('by heart'), involves an almost purely acoustic memory-mechanism. The learning and mental (or technical) reproduction of a difficult musical work ('by memory') requires considerable motor (finger, pedal, voice, bow) memory. (The often observed unconscious humming or 'playing' on the school desk by musical children despite their attention to the teacher may represent an objective proof of unconscious activity of the musical memory-mechanism.) Visual memory is limited to faint imagination of the instrument and its handling; only in rather unmusical persons does the visualization of the musical script play a rôle. In persons with absolute pitch a very strong association exists between pitch (heard as well as remembered), its mechanical production, and its musical name. Good musicians possess astonishingly fine memory-hearing for complicated

melodies and harmonies. Musical memory accordingly may be considered as a Gestalt-process, based upon the integration of facilitated reflexes and memory-records into an extremely complex engram. (f) The most astonishing part in this recording is the time-element. It is of extreme accuracy from the fastest finger-passages and finest coördination to the temporal structure of a concerto or a symphony. Even the delicate switch-mechanisms at the points of deviation (within a melody, for repetitions of a part and for the main architecture of a movement) can work without errors, as well in conscious as in unconscious memory-playing. All this is possible for concert-musicians without the help of a prompter. In all these respects the musical memory-mechanism is by far superior to all other memory-mechanisms in respect to complexity, accuracy, reproductivity, and reliability.

Conclusions. On the basis of all these facts and Penfield's doctrine of memory and consciousness, one may postulate that the memory-cortex is capable of sustaining highly complex, associative and integrative, well-timed activity which remains unconscious if disconnected from the centrencephalon, but becomes conscious when such a connection is established. A close analogy exists between reflex activities, which may either go on unconsciously (in relative disconnection from the centrencephalon) or consciously (in active contact with that center), and cortical activity, which may either take place unconsciously or consciously (also depending upon relative isolation from, or active contact with, the centrencephalon). This analogy is of particular interest in view of the fact that the connections of the reflex arcs and those of the memory-pathways with the centrencephalon are opposite ones, *i.e.* reticulo-peripheral and reticulo-cortical respectively.

Department of Physiology
College of Medicine
University of Illinois

ALBERT BACHEM

SENDERS ON RESPONSE-SEQUENCES

The recent article by Senders¹ on the non-independence of successive responses deserves comment. Although sequential dependencies for re-

¹ V. L. Senders, Further analysis of response-sequences in the setting of a psychophysical experiment, this JOURNAL, 66, 1953, 215-228.

sponses made to threshold values of the stimulus have been demonstrated,² it does not follow in all cases that "the basic assumption of almost all our statistical treatments is in question,"³ nor that the theoretical concepts of

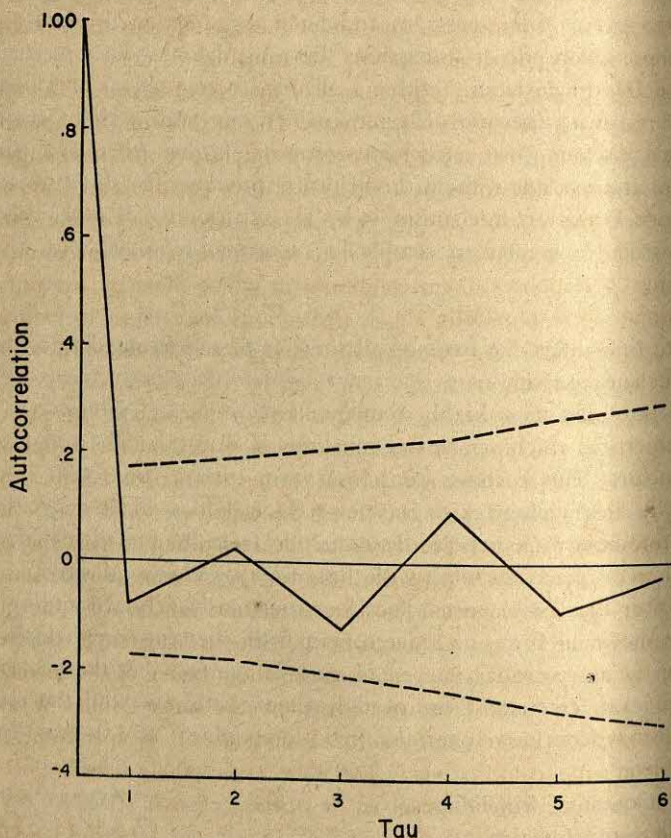


FIG. 1. AUTOCORRELATION FUNCTION OF RESPONSES AT THRESHOLD

absolute and differential limens are seriously undermined. There is a line of demarcation between *guessing habits* and *psychophysical judgments*.

When *O* is required to respond continuously to threshold stimuli or to discriminate between two equivalent stimuli over a long series of trials,

² J. P. Flynn, Lack of randomness in sequences of auditory differential threshold data, *Amer. Psychol.*, 3, 1948, 254; G. H. Collier, An investigation of the independence of successive responses for behavior at the visual threshold, *Amer. Psychol.*, 6, 1951, 278.

³ Senders, *op. cit.*, 228.

does he have any other alternative than to guess? This was the condition facing Senders' *Os*. In a psychophysical experiment, however, there are a large number of supraliminal as well as subliminal presentations. A judgmental *set*, not a guessing game, develops (especially when *O* is allowed three categories of response), and under these conditions the assumption of response-independence at threshold is reasonable.

Seven *Os*, drawn from the personnel of the Naval Research Laboratory, served in a brief test of this hypothesis. The method of constant stimuli was used, *Os* being instructed to report on intensive differences between standard and variable tones of 1000 cycles. Five equally-spaced intensities were used as the variable stimuli—two above and two below the standard stimulus, and one identical. Each *O* made a total of 100 discriminations including 20 responses under conditions in which the standard and variable stimuli were physically equal. Precautions were taken to maintain a constant time-interval between the successive presentations and to counter-balance the time-error.

The threshold-responses were analyzed as a function of the preceding supra-threshold responses by the autocorrelation method for values of τ from 0 to 6. This analysis was subject to one restriction, *i.e.* that the response to be correlated with the threshold-response should not be a previous threshold-response. Responses to the real differences in the paired stimuli were correct 98.9% of the time for the extreme comparisons and 89.3% for the less extreme. The autocorrelation function for the pooled data is shown in Fig. 1. Although periodicities are suggested, the magnitude of the autocorrelations is extremely small and none of the correlations is significant. *Os* treated individually show the same results. While no single study conclusively proves the hypothesis of independence, these results lend support to that view. The writer trusts, along with Dr. Senders, that this question will be "resolved by further research."⁴

Ohio State University

JACK E. CONKLIN

TWENTY-FIFTH ANNUAL MEETING OF THE EASTERN PSYCHOLOGICAL ASSOCIATION

The Eastern Psychological Association met on April 9 and 10 at the Hotel New Yorker in New York City. A total of 1,857 persons registered for the meetings. Of these 1,034 were members of the Association, 571 were guests, and 252 were new members who joined the Association at

⁴ Senders, *op. cit.*, 228.

the meetings. The present active membership of the Association totals 2,487.

The program consisted of 156 scientific papers (presented in 21 sessions), 5 symposia, and 3 special meetings. The 21 sessions were concerned with the following topics: brain mechanisms, animal learning (3 sessions), visual perception (2 sessions), general experimental (2 sessions), perception and personality, evaluation and measurement, motivation, projective techniques, industrial and military, social psychology, sensory processes (2 sessions), human physiological, clinical and social (2 sessions), personality and learning, and human engineering. The symposia were as follows: "The criterion dilemma in human engineering," "Problems and difficulties in evaluating patient progress in therapy," "Personality evaluation for industrial purposes," "Psychodynamic patterns in the pedophile sex offender," and "Problems in the use of projective techniques in research." Special meetings included a meeting of Psi Chi, and meetings bearing the titles "Public opinion and foreign policy decisions: informal discussion of research problems," and "The psychologist in private practice and the public."

Dr. Harold Schlosberg presented the annual presidential address, "Fatigue, effort and work output." New officers elected were B. F. Skinner, President; Carl Pfaffmann and Richard L. Solomon, Board of Directors.

The 1955 meetings of the Association will be held at the Benjamin Franklin Hotel in Philadelphia on April 15 and 16.

University of Delaware

GORHAM LANE

TWENTY-FOURTH ANNUAL MEETING OF THE ROCKY MOUNTAIN BRANCH OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

The Rocky Mountain Branch of the American Psychological Association held its twenty-fourth annual meeting with the Psychology Section of the Colorado-Wyoming Academy of Science on April 30 and May 1, 1954, at the University of Wyoming, Laramie, Wyoming. About 120 members and guests attended the sessions. Twenty-two papers were read at three sessions. The papers were primarily in areas of clinical and educational psychology and in group dynamics. At a luncheon meeting, the twenty-fifth anniversary of Psi Chi in the Rocky Mountain area was celebrated.

A symposium was held on "Training and function of school psychologists" under the chairmanship of Victor C. Raimy of the University of Colorado. The participants were Owen Caskey, Veterans Administration, Denver; F. Allen Murphy, Public Schools, Denver; and Philip Buck of Gunnison, Colorado. Robert H. Bruce of the University of Wyoming chaired a symposium on "Poor performance as a function of inadequate learning." The participants were: Kenneth B. Ashcraft, Gates Rubber Company (Industrial); Lawrence E. Dameron, University of Denver (Social); E. Ellis Graham, University of Denver (Child); Richard Jessor, University of Colorado (Clinical); Arno H. Luker, Colorado State College of Education (Educational).

Hugh B. McFadden, President of the Branch, served as Chairman of the business meeting. It was voted to continue publication of the minutes of the meeting and the program as in previous years. Further attempts will be made to broaden participation of psychologists in neighboring states in branch activities.

The officers of the Branch elected for the year 1954-55 are: president, Lawrence S. Rogers, Veterans Administration, Denver; and secretary, Wilson J. Walthall, Jr., University of Wyoming. Virginia M. Brown of the Lowry Air Force Base continues as treasurer.

Denver, Colorado

LAWRENCE S. ROGERS

FIFTIETH ANNUAL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The fiftieth annual meeting of the Society of Experimental Psychologists, the Golden Jubilee, was held at the Statler Club, Cornell University, on April 1 and 2, 1954, in honor of the founder of the Society, Edward Bradford Titchener. James J. Gibson, as Chairman, presided at the business meeting, during the morning of April 1 and at the subsequent scientific sessions.

The following members were present at the meetings: Bray, Boring, Brown, Fernberger, Geldard, Gibson, Graham, Grant, Harlow, Hunter, Irwin, Keller, Kennedy, Landis, Licklider, Miles, G. A. Miller, N. E. Miller, Morgan, Mueller, Nafe, Pfaffman, Riggs, Schlosberg, Smith, Underwood, Volkmann, and Wendt.

At the business meeting, W. K. Estes, R. M. Gagne and W. R. Garner were elected to membership, bringing the total of active members to 52.

Three scientific sessions were held at which members reported informally

on research recently completed or in progress. The Department of Psychology in Morrill Hall was open for visits during the afternoon of April 1.

Following the dinner on April 1, the Warren Medal for 1954 was awarded to Neal E. Miller "for his distinguished contributions to the scientific investigation of the relationship between learning and emotional behavior, leading to an increased understanding of the development and fixation of emotional attitudes."

It was voted to accept the invitation of Brown University to hold the next meeting in the spring of 1955. Harold Schlosberg was elected Chairman of the Society for the coming year.

Santa Monica, California

JOHN L. KENNEDY

TWENTY-SIXTH ANNUAL MEETING OF THE MIDWESTERN PSYCHOLOGICAL ASSOCIATION

The Midwestern Psychological Association met at The Neil House, Columbus, Ohio on April 30 and May 1, 1954, under the sponsorship of the Ohio State University. Delos D. Wickens served as chairman of local arrangements. The total registered attendance was 940. The program committee, under the chairmanship of Jay L. Otis, presented a program of 115 papers and 8 symposia. Particularly prominent were papers on probability learning, anxiety, and group performance, in addition to the traditional topics of learning, vision, social psychology, and personality. Few clinical papers and few papers dealing with applied psychology were received. The topics for symposia were as follows: psychology in general education, psychology in medical education, problems of human engineering, theory of psychological measurement, recent developments in semantic measurement, the automobile driver, construct validity in psychodiagnostics, and "Can the psychology of learning be applied?" Judson S. Brown of the State University of Iowa delivered the presidential address, "Pleasure-seeking behavior and the drive-reduction hypothesis."

Paul E. Meehl, University of Minnesota, was elected president for 1954-55, and Julian B. Rotter, Ohio State University, was elected to the executive council for a three-year term. Election of 362 new members brought the membership of the Association to 1526, a net increase of

17 per cent. The 1955 meeting will be held on April 29-30 at the Sherman Hotel, Chicago.

University of Illinois

LEE J. CRONBACH

FORTY-SIXTH ANNUAL MEETING OF THE SOUTHERN SOCIETY FOR PHILOSOPHY AND PSYCHOLOGY

The forty-sixth annual meeting of the Southern Society for Philosophy and Psychology was held at the Biltmore Hotel, Atlanta, Georgia, April 15-17, 1954, with the Philosophy and Psychology Departments of Emory University as hosts. Approximately 275 members and guests were present. Twenty-six papers in philosophy and 51 papers in psychology were presented. At the joint session a symposium was held on "Language and the unity of science." James G. Miller and Roger Buck were the principle speakers, representing philosophy and psychology respectively.

Following the annual banquet, Karl M. Dallenbach delivered the presidential address entitled "Phrenology versus psychoanalysis."

At the business meeting, one member was transferred from the status of associate to full membership; and 10 new associate and 27 new members were elected to the Society. The new officers elected were: Charles A. Baylis, President; Joseph E. Moore, Secretary; Nicholas Hobbs, Treasurer; Lewis Hahn, Council Member for Philosophy; and Edward E. Cureton, Council Member for Psychology.

The 1955 annual meeting will be held either in Biloxi, Mississippi, or in New Orleans, Louisiana, under the auspices of Tulane University during the week end of the Easter vacation.

University of Alabama

OLIVER L. LACEY

THE 1954 MEETING OF THE NATIONAL ACADEMY OF SCIENCES

The annual meeting of the National Academy of Sciences was held in Washington on April 26-28, 1954. Of the 24 members of the Section of Psychology, 11 were present—Beach, Boring, Gesell, Graham, Lindsley, Miles, Pillsbury, Richter, Skinner, Stevens, and Woodworth. J. P. Guilford was elected to membership.

Frank A. Beach, in collaboration with Allen Goldstein and George Jacoby, read a paper on the "Effects of electroconvulsive shock on mating

behavior in male rats," showing how shock alters sexual activity temporarily and may increase the capacity for mating performance permanently or at least for long periods.

In the exhibit of devices, instruments, and apparatus, B. F. Skinner, assisted by Charles B. Ferster, demonstrated the recording of the behavior of a pigeon, which, for each of three different light signals, exhibited the response pattern for one of three different schedules of reward: reinforcement at fixed intervals, reinforcement at a fixed ratio of intervals, and variable reinforcement in accordance with a special schedule.

Harvard University

EDWIN G. BORING

THE 1954 MEETING OF THE AMERICAN PHILOSOPHICAL SOCIETY

The American Philosophical Society held its annual meeting at the House of the Society in Philadelphia on April 22-24, 1954. Of the 14 psychologist-members, Boring, Hovland, Hunter, and Miles were present.

Wilder Penfield read a paper on "Some observations on the functional organization of the human brain," in which he described the nature of the imaginal phenomena experienced by conscious subjects when the ideational cortical areas for the different senses are electrically stimulated on the exposed cortex. The Penrose Lecture by Vannevar Bush, given under the title "Scientific motivation," dealt with methods of storing and making available the vast quantities of knowledge which are being created in the modern scientific age and the maintenance of scientific effectiveness under this increasing burden of responsibility.

Harvard University

EDWIN G. BORING

BOOK REVIEWS

Edited M. E. BITTERMAN, University of Texas

Social Psychology. By SOLOMON E. ASCH. New York, Prentice-Hall, 1952. Pp. xvi, 646.

Social psychology is still vague. There is no standard pattern for a course on the subject, and there is no fixed list of topics with which every textbook must deal. The writer is still free to present as social psychology what he thinks is "really important for the understanding of man." For this reason the new books on social psychology are particularly interesting. Our publishers insist that we write textbooks; but only in a textbook on social psychology can we write as we please. The past few years have presented us, in the guise of textbooks on social psychology, with a wealth of major systematic contributions to the understanding of man. Krech and Crutchfield, Newcomb, Sherif, Hartley and Hartley, Queener, and Sargeant—to mention but a few—represent serious attempts to grapple with the problem of understanding man as a member of society. Asch's new book is certainly to be ranked with the best of these. We regret that it had to be presented as a textbook—for it is clearly not a textbook—but we like to think that the better students of psychology will find in its mature and thoughtful consideration of the many-faceted problems of social behavior a greater challenge than is encountered in the usual text. Asch's book is a serious attempt at psychological systematization, governed by a point of view that is explicitly stated and defended, and written in a style that assumes a literate and thoughtful reader. Our educational efforts would be more successful, and teaching would be a more exhilarating occupation, if we had more books of this calibre towards which we could steer our students. A course centered about this book might leave the student with fewer 'objective facts' than he would glean from a more conventional course, but the student would emerge a better educated man.

The style of the book is impressive. Although Asch makes no reference to Brentano, the reviewer finds the comparison irresistible. This is a book in the tradition of *Die Psychologie vom empirischen Standpunkte*. It has the same graceful prose, the same fine feeling for the correct expression and the apt metaphor, and the same wealth of reference (not always explicitly identified) to sources in literature, the arts, philosophy, and the other sciences. Asch as a student of man has steeped himself not merely in the literature about man but in the cultural products through which the essential nature of man has found expression.

Again like Brentano, Asch's approach is empirical but not primarily experimental. For Brentano the essential facts of experience need no experimental confirmation, if one can apprehend them clearly and think about them logically. Psychological experiment can help to sharpen one's observation but, in a sense, its main function is not to reveal new truth but to demonstrate to others truths that in the last analysis must be tested by intuition. The bulk of Brentano's *Psychologie* is consequently a careful balancing of *pros* and *cons* in the attempt to clarify intuition.

The bulk of Asch's book is similarly a balancing of *pros* and *cons*. He is manifestly appealing to us as social beings to observe our own experience, to reflect on it, and to judge whether or not it can be explained satisfactorily in any other way than that which he proposes. He enlarges our experience by reminding us of things we may have forgotten and by reporting observations in cultures other than our own. Here and there he invokes the (spurious?) criterion of universality, but the final criterion is always the intelligibility of the finding in the light of our own experience. Experiments and other 'empirical' studies are quoted, some at considerable length, but, as in Brentano, they usually appear as illustrative demonstrations rather than as sources of new truth.

The resemblance to Brentano might be extended to the content of Asch's doctrine of man, but that would be an unnecessary forcing of the parallel. Through the tradition of Gestalt theory Asch has undoubtedly been influenced by Act Psychology, and his rejection of the atomistic, reductionistic bias is certainly in keeping with Brentano's emphasis. Much has happened since 1874, however, and Asch's thinking deserves to be considered in the context of the mid-twentieth century.

What Asch gives us is a mid-twentieth century doctrine of man, a doctrine for people who are seeking a workable democratic alternative to communism and fascism, who are trying to find a place for individuality in a society that requires increasing amounts of social control, whose constant exposure to propaganda has made them wonder whether anything can be really believed, for whom the reported facts of cultural diversity appear as threats to traditional standards of ethical conduct. Has the psychologist of today any message that will help to restore sanity and a measure of optimism to a world that is confused? Asch is convinced that psychology has such a message, but that before it can be made intelligible many of the assumptions of current psychology must be challenged and many of the facts of human behavior must be scrutinized anew. Two convictions run through the book, one that the currently popular systematizations of psychology are inadequate, the other that a radical reorientation of our approach, with man-in-society as our starting point, will enable us to develop a system that is worthy of our subject-matter. The recurrent pattern of the book is, accordingly: preliminary statement of the problem, demonstration of the inadequacy of the common approaches, closer examination of the evidence, restatement of the problem in its larger context, affirmation of a sounder type of theory. There will be readers who will disagree with Asch's critique of other theories, with his evaluation of the evidence, and with the type of theory to which he consistently returns. None will fail to recognize the importance of the problems as he states them, nor fail to respect the courage and admire the skill with which he attacks them.

Asch's doctrine is, broadly speaking, a Gestalt theory of man as a social being. It leans heavily on Gestalt investigations of perception, motivation, thinking, and development, and it represents in many way an elaboration of ideas which Wertheimer, Köhler, and Duncker advanced in their discussions of value theory. No other Gestalt theorist, however, has attempted to work out the full implications of the theory for social psychology. Lewin made brilliant forays into various fields of social behavior, but never integrated his findings into a system. Krech and Crutchfield reflected a Gestalt approach but, as authors of a textbook, evidently felt pressed to incorporate a great deal of what is conventionally known as social psy-

chology and to exclude evidence not found in 'the literature.' Asch has been bolder. He challenges us to look, with the eye of a psychologist, at all the complex problems of man-in-society, to survey the phenomena in all their dimensions, and to let the phenomena themselves dictate the theory.

The Gestalt approach has demonstrated its value in the psychology of perception. Properties of organized wholes are no longer rejected as invalid because no simple counterparts in stimulus or receptor can be found. The recognition of their psychological reality has led to the discovery of new principles of organismic functioning, and has assisted psychologists in bridging the gap between psychological and physical theory. Can the Gestalt approach be extended to the study of social phenomena, with the consequent implied hope that, through psychology, the social and natural sciences can eventually be unified? Asch seems to be buoyed by the belief that the realm of Natural Law includes the social behavior of man, that Social Law and Natural Law are in the last analysis one. If we accept Asch's faith, we may hope eventually to solve in a rational way the problems of social living.

A Gestalt theory of social behavior, like a Gestalt theory of perception, begins with phenomenology. It attempts to 'bracket' all biases and look at phenomena as they are, proceeding therefrom to the establishment of correlations and the construction of theory. Although Asch does not use the term, he is in this sense a phenomenologist. The data of social psychology are just like the data of any other kind of psychology: things, events, spatial and temporal relationships, selves, anxieties, aspirations, and the multitude of unclear cognitive structures about which we are oriented. All must be accepted initially at face-value, and their implications examined, before they can be disposed of as products of 'association by frequency' or 'the law of effect' or 'ego defense.' The world that people live in contains goals and values, theories and beliefs, just as it contains pains and hungers, frustrations and doubts. Asch, in formulating the problems of social psychology, attempts to keep all the phenomena in perspective. As a scientist, he is challenged just as much by man's achievements as by man's failures.

Asch's *bêtes noires* are, of course, the thought-patterns of the nineteenth century. According to nineteenth century physics we could explain by reducing to atoms and their interrelationships; according to Darwinian biology we could find explanations by searching for origins and tracing the process of development. Both approaches are reductionistic. Both are reflected in modern stimulus-response theory and in modern Freudian theory. Any theory, Asch would argue, that is based on simple associationism and hedonism, is so restrictive as to render impossible a psychology that can deal competently with the full richness and variety of human experience and behavior. Granted that there may be alternative, equally consistent, ways of viewing man, no psychology is adequate if within its model the distinctively human characteristics of man cannot be encompassed, his rationality as well as his irrationality, his creativity as well as his stupidity, his loyalties, his ambitions, his ethical ideals and, above all, his capacity for shared experience. The psychologist, above all, should be curious about man at his best. In this respect, Asch feels, both the behavioristic and the psychoanalytic approaches are incapable of leading us to a full understanding of man.

Whether or not Asch has made a proper case against his opponents will be disputed. Any polemic is open to the 'straw-man' criticism, and Asch's is no excep-

tion. It is much easier to dispose of an enemy if one has disarmed him first, and many Freudian and Hullian disciples will wish that their champions had had a chance to fight back. McDougall's followers will feel the same way. What must be remembered, however, is that this is an argumentative book. Although Asch quotes liberally from representatives of alternative views, he is dealing essentially, not with particular theorists, but with types of theory. Asch has attempted to shear away personalities and other irrelevancies and to expose the naked issues to the clear light of logic. Whether or not we finally agree with his reasoning, we now have a clearer picture of the issues.

From the point of view of this reviewer, however, the real question centers about our view of psychology as science. Many readers will find the book an interesting, thoughtful, essentially 'philosophical'—but not 'scientific'—disquisition on human nature. They will seek in vain for nice, neat principles in terms of which all behavior can be explained, predicted, and controlled. To oversimplify the situation, one might say that our psychological theorizing seems to be either model-oriented or problem-oriented. The model-builders begin with a model—it may be sociological, psychological, neuro-physiological, physical, or mathematical—and try to incorporate all possible phenomena into the model. Phenomena that do not fit are rejected or postponed. Titchener and Watson were in this sense model-builders; so were Freud and Hull. The problem-oriented psychologist is less likely to worry at first about his model. He begins with a massive question, looks at the relevant phenomena, and then tries to construct a model that will account for all the phenomena. If the model does not fit the phenomena, he constructs another model. Lewin was one of these. Finding Euclidean geometry unsatisfactory, he tried topology; and finding that topology could not represent the dimension of force, he introduced vectors. The result was still unsatisfactory, for it did not include the temporal variable in behavior. If he had lived longer, he would certainly have revised his model still further, but the revision would have been made in response to the dictates of the phenomena, not to the requirements of the model. The model-oriented theorists give us a neat theory that will account for some of the phenomena; the problem-oriented theorists give us a ragged theory, that attempts to encompass all the phenomena, but that may be discarded as soon as the phenomena point in a different direction. Both types of theorizing have proved their usefulness in the history of science.

Asch clearly belongs to the problem-oriented group. His massive question—a very old one—is: How can people live together in society so as to achieve the good life? The phenomena he looks at are not the acquisitive tendencies of squirrels, the salivary conditioning of dogs, or the college sophomore's learning of nonsense-syllables. He looks at the most 'human' aspects of man's behavior, the goals towards which behavior is directed rather than the biological needs that may initiate it, the values, beliefs, and prejudices that are found in organized society rather than the hypothetical learning processes whereby they may have been acquired, man's occasional resistance to conformity as well as his tendency to follow the herd. He insists that we study the complexities as well as the simplicities of life, and that we try to see these complexities in terms of 'what they mean to the individual.' This is a much more difficult task than most social psychologists have been willing to face. To systematize this kind of psychology is a major undertaking.

Many readers will be of the opinion that Asch has failed, and this reviewer tends

to share that opinion. But it is a brilliant failure. Asch has honestly and courageously faced the really important problems of social psychology, he has examined the conventional solutions and found them wanting, but he has not given us a new theory that fully illuminates the facts. The fault, however, lies not in the author but in a science that has been more concerned with models than it has with phenomena. Asch has not given us an adequate science of human behavior for the simple reason that in our psychological researches we have not been looking at behavior in its truly human aspects, and that consequently we lack the factual resources for good theorizing. Asch has, however, reformulated the problems, pointed to phenomena that we ought to be considering, and indicated the proper line of inquiry. Sir Francis Bacon is reported to have said: *Prudens quaestio dimidium scientiae*. In this sense, Asch has substantially advanced the science of psychology. If to be a scientist is to be persistently curious about phenomena, to continue looking, wondering, and theorizing, then Asch's book will rank as one of the milestones in the history of scientific social psychology.

Cornell University

ROBERT B. MACLEOD

A Behavior System: An Introduction to Behavior Theory Concerning the Individual Organism. By CLARK L. HULL. New Haven, Yale University Press, 1952. Pp. ix, 372.

It is symbolic of Hull's life that this final volume, which is important in rounding out the presentation of his learning theory, was completed only three months before he died, and was completed despite the fact that, in his last few years, his health permitted him to work only a few hours each day. As one may judge from his simply-told story in the *History of Psychology in Autobiography* (vol. IV), there have been relatively few psychologists who have had to fight so hard against early cultural disadvantages and against physical handicaps. It is not surprising that those who were close to Hull saw certain heroic qualities in his work and were stirred to unusual loyalty to him.

It is fortunate for psychology that Hull was able to finish this volume, which was not the last one he had hoped to write. His plan called for still another volume to apply his theory to social or group behavior, but this further volume would have been parallel in many ways to the present one and less important in rounding out the presentation of Hull's basic thinking. The present volume, on the other hand, is fairly indispensable. It clarifies Hull's goal and methodological concepts as a learning theorist; it presents his final conclusions on various details of interpretation; it gives the basis for discussing the adequacy and degree of empirical support of his hypotheses; and, finally, it puts us into a better position to appreciate the general character of Hull's contribution.

Hull's goal. This book, as everybody would have anticipated, illustrates very clearly Hull's general theoretical orientation. It illustrates as well, however, the more specific form in which Hull envisaged his goal. Hull has made this point in the autobiographical paper mentioned above: "... I came to the definite conclusion around 1930 that psychology is a true natural science; that its primary laws are expressible quantitatively by means of a moderate number of ordinary equations; that all the complex behavior of single individuals will ultimately be derivable

as secondary laws from (1) these primary laws together with (2) the conditions under which behavior occurs; and that all . . . strictly social behavior . . . may similarly be derived . . . from the same primary equations . . . the task of psychologists obviously is that of laying bare these laws as quickly and accurately as possible, particularly the primary laws. This belief . . . has determined the most of my scientific activities ever since, and the longer I live the more convinced I am of its general soundness" (p. 155). Phrased in this way, Hull's goal was still one with which most psychologists would sympathize. They might doubt that the full range of human activity will ever be derived from a "moderate number" of equations; but they still would agree otherwise.

Hull's belief as to the best strategy for reaching his goal. We are familiar nowadays with the concept that research ought to be guided by an explicit methodology. In much the same way, to think effectively about Hull's work in learning theory we must understand explicitly the *methodology of theorizing* on which Hull pinned his hopes. There are three features of this method which should be noted. (1) As the source of his basic principles, Hull preferred to consider, not the whole range of psychological research, but experiments within only a very restricted area. He felt that the essential phenomena of learning and performance were apt to show themselves in an objective and understandable fashion in classical and instrumental conditioning more than in any other situation. Even within this area, to get his basic postulates he disregarded experiments on differentiation and inconsistent reinforcement. (2) To secure tentative quantitative formulations from the studies within this miniature area, his procedure was that of selecting, typically, only one or two experiments from the body of expert studies. He fitted equations to the particular values reported in these selected studies and (especially in this final volume) included the constants from these equations in his postulates to give the quantitative character which he felt they should have. (3) To test the worth of the quantitative laws thus secured, Hull derived from these postulates various theorems with regard to effects that should be seen in discriminative learning, maze learning, problem-solving, and several other types of experimental situation. He then sought to compare these theorems with relevant experimental data.

When I reviewed Hull's previous book, *The Essentials of Behavior* (1951) (this JOURNAL, 65, 1952, 478-491), I noted that Hull was developing such a 'miniature-area system' and I expressed my uncertainty as to what Hull intended to get from it. As I said, "Since there is a good deal of talk about such 'miniature systems' in psychology, it is important for us to clarify our ideas about what they accomplish. There are several possibilities. . . . One is that research within such a miniature area can establish the principles from which all of the more complex phenomena outside that area can be deduced. Sometimes it seems that Hull takes this position" (p. 489). Now we can see that this last sentence should have read, not "Sometimes . . .," but "Basically, this is Hull's faith."

The content of this book as an application of this strategy. In the two preceding books in his series—*Principles of Behavior* (1943) and *Essentials of Behavior*—Hull sought primarily to work out the first methodological task of the three outlined above. That is, he sought primarily to derive a set of basic principles or hypotheses, from classical and instrumental conditioning experiments. He gave some illustrations of the possible applications of these principles to other fields

of learning, but he has emphasized in subsequent writings that he was not then attempting to demonstrate the general validity of these basic postulates. In the *Principles*, Hull also attempted to get his laws into rough quantitative form. This work was limited, however, to attempts to indicate the *type* of curve involved in each case, and there was no attempt to find exact values for the constants used in the equations. In the *Essentials*, a few definitely new concepts were added—especially the concept of 'stimulus intensity dynamism' and the concept of a scaling method of expressing habit strength.

In the present volume, a minor part of the discussion continues the effort of the two preceding volumes, with simple non-competition situations being used to decipher the hypotheses to be taken as 'primary laws.' Eleven of Hull's 17 main postulates are now presented with specific numerical values as constants. The figures which previously seemed to be only illustrative are now incorporated, without any additional evidence of their representativeness, within the primary laws. Most of the space in the present volume is, however, devoted to the deduction of a long series of 'theorems' regarding somewhat more complex learning situations and to surveying (sometimes in detail, sometimes in a very sketchy manner) the experimental evidence bearing on about 65% of these theorems. Many of these deductions follow the approach developed by Spence with reference to discriminative learning: positive reinforcements build up positive habits which tend to generalize along the afferent continuum; non-reinforcements tend to build up habits of non-responding and these likewise tend to generalize; gradually an excess of reaction potential sufficient for consistently correct responses is developed. Many of the other deductions depend on a new approach to the goal-gradient principle, which is derived from the stimulus-intensity dynamism and the gradient of the 'molar stimulus trace.' As a whole, the present volume is very difficult. Unless a reader has a lot of experience with Hull's previous writing, and unless he is willing also to take a lot of time to figure out why Hull used this calculation or that, he will simply have to take on faith a great part of the discussion.

I think there is no question but that the present volume, more clearly than the earlier ones, indicates that Hull's theory is basically a fairly atomistic reinforcement theory. For example, even the development of skill is interpreted as a matter of the development of a large number of different habits, each related to different 'strands of muscular contractions.' The proposal is that, since reinforcements follow awkward movements as well as better-adapted movements, the muscular contractions involved in occasional awkward movements acquire some habit-strength too, but eventually lose out to better-adapted muscular movements because they generally have a greater temporal separation from the reinforcement. Blind alleys are eliminated for the same reason.

As mentioned above, the majority of Hull's postulates give in their present form the impression of very refined quantitative work. For example, Postulate V describes in the following way the strength of the hunger drive in rats as a consequence of different hours of food deprivation:

$$D = (37.824 \times 10^{-27.496} \times 1/h + 4.001) \times (1 - .00001045^{2.486})$$

Often, as in this case, values are given as though they were significant to four or five figures. Actually, however, there is no evidence to indicate that the quantitative

values are significant in anything more than the very roughest sense. In the first place, there is nothing in their origins to indicate that the specific values are worth noting. Thus, as I mentioned in the review of the *Essentials*, Hull's new procedure for absolute scaling of habit strength was used by two of his students with different outcomes, and was used by Zeaman with still other results. Hull merely adopted *one* of the resulting equations and left unmentioned the fact that there were other results in other studies. Often, furthermore, the source studies are not highly reliable and ought not to have been treated as giving strong support for one curve rather than some other. Finally, there really is nothing in the present volume to demonstrate that the precise quantitative values are essential to the deduction of the various theorems. Hull could have taken a wide range of values and still ended with the same general predictions. Thus, as far as the present book gives evidence, I think that all of his quantitative expressions might just as well have been replaced by statements to the effect that a given variable increases or decreases with changes in some other variable.

The value of Hull's principles. In his last chapter, when he looked back over his book, Hull decided that 123 of his 178 theorems had been compared with relevant evidence, that 106 of the 123 appeared to be definitely validated, most of the rest had some support, and only one was definitely challenged by available evidence. Such material shows that the general features of Hull's principles have some value. The evidence is not, however, extensive. For example, to show that his concepts could account for discrimination-reversal learning, Hull considered two rather small studies which used conditions whose influence could be harmonized more or less with his postulates, but he did not mention the more extensive experiments by Harlow which his theory hardly could have accounted for. To explain the effects seen in Crespi's experiment on changes in amount of reward, Hull, assumed a rapid emotional learning based on frustration, but he did not examine the question of whether, from his standpoint, the conditions should have yielded any frustration. To account for transposable habits in some discrimination experiments, he considered the possibility that, in experiments with simultaneous presentation of the positive and negative stimuli, the animal learns to respond to the "degree and nature of the afferent shift from the trace of the preceding stimulus to the stimulus which follows"; but he did not indicate how this proposal might have been harmonized with such further effects as those obtained by Saldanha and Bitterman in 1951. He deduced that, where animals have no pre-existing preferences for one alternative rather than the other, as in usual discrimination experiments, correct choices should begin to prevail (with large groups) after the first trial or two, but he did not note the findings of discrimination experiments which do not support this conclusion.

Such difficulties could be multiplied, but, as I followed Hull's successive books, it finally dawned on me that it was pointless and is pointless to indicate such discordant evidence. Originally, I thought that Hull described his main aim when he said that "... empirical observations are regarded as primary, and wherever a generalization really conflicts with observation the generalization must always give way" (*Principles*, p. 12). With this final volume, however, I come to see that Hull over-estimated his interest in empirical details and under-estimated, at least in such statements, his interest in developing a comprehensive theory. His heart was in the

latter task, in much the way that James B. Conant has described in his discussion of scientific theorizing in his little book *On Understanding Science*. As Conant says: "... in complex affairs of science, one is concerned with trying to account for a variety of facts and with welding them into a conceptual scheme ... a conceptual scheme is never discarded merely because of a few stubborn facts with which it cannot be reconciled ... it takes a new conceptual scheme to cause the abandonment of an old one ..." (pp. 84, 89). It is the sort of thing that Conant illustrates by reference to the controversy over phlogiston between Lavoisier and Priestley: "Not until long after Lavoisier's execution in 1794 was the relationship between the two gases straightened out. Therefore, Lavoisier was never able to respond to the most weighty of Priestley's arguments against his doctrine. He merely ignored the alleged facts, much as Priestley ignored the unexplained gain in weight or calcination. Each undoubtedly believed that some way would be found around the difficulty in question. ... So proceeds the course of science" (p. 96).

About all that we can say at this point, therefore, is that Hull's theory represents an earnest attempt to develop a comprehensive account of learning and that it has at least the value of dealing explicitly with a wider range of phenomena than perhaps any alternative interpretation. It cannot by any means handle all of the difficulties which a skeptic could have presented to Hull, but, as Conant says, "Sometimes it turns out that difficulties with a concept or conceptual scheme are wisely ignored, sometimes unwisely" (p. 96). It remains to be seen which will prove to be the case with Hull's theory. In line with Conant's interpretation, perhaps what we need most is the further development of alternative lines of thought so that we can have the means of comparing Hull's theory with equally inclusive or more inclusive alternatives.

University of Oregon

ROBERT LEEPER

Human behavior from the transactional point of view. F. P. KILPATRICK, Editor. Hanover, Institute for Associated Research, 1952. Pp. vii, 259.

For several years now the 'Ames demonstrations,' under various auspices, have won the attention of workers in widely different fields—psychology, philosophy, education, and art. The demonstrations themselves, the experiments based on the demonstrations, and the theory generated by the demonstrations and experiments, are generally accepted as proving the necessity for a fundamental reorientation to the problems of perception. The volume under review is largely a compilation of experimental investigations and theoretical papers already published or about to be published. In a sense, then, it represents a certain definite stage, though by no means a final one, in the exposition of the Ames-Cantril point of view.

The monograph is divided into three main parts. Part I deals with the "demonstrations and theory." The chief demonstrations, with which the reader is probably already familiar, are the 'chair,' the 'distorted room,' the expanding and contracting balloons, the 'card' setup, the rotating trapezoidal window, 'radial' movement, and the perceptual effects produced by the wearing of aniseikonic glasses. This section closes with a previously unpublished paper by Kilpatrick on a "statement of theory." Part II presents "some representative research papers." Four of the seven papers of this section have not yet been published separately. There is

a study by G. H. Smith on perceptual changes associated with distance judgments of "friendly" and "unfriendly" faces; a study by Hastings on the relationship between "visual perception" and "level of personal security"; a paper by Kilpatrick on "assumptions and perceptions: three experiments" which is concerned with the influence of "assumptions" on several perceptual effects derived from the rotating trapezoidal window; and a fourth paper by Ittelson and Kilpatrick which deals with the same window. Part III considers "some implications of the demonstrations and theory." It consists of a group of papers in which the nature of scientific research, social perception, and the 'constancies' are discussed from the Ames-Cantril point of view.

Since the "transactional approach" to perception can be assumed to have reached a point of maturity in the monograph under review, it may be worthwhile to point out the basic theoretical postulates. According to Kilpatrick, "a given physiological stimulus-pattern may be related to (produced by) an infinity of different external conditions. Consequently, the particular external configuration perceived cannot be derived from the stimulus-pattern alone; only a class of external configurations may thus be derived" (p. 3). In examining the implications of this statement Kilpatrick continues at another point in his exposition: "But we never see an infinity of configurations; we see just one. This means, of course, that perception cannot be 'due to' the physiological stimulus-pattern; some physiological stimulus probably is necessary, but it is not sufficient. There must be, in addition, some basis for the organism's 'choosing' one from among the infinity of external conditions to which the pattern might be related" (p. 88). The 'choice' which the individual makes is a function of his past experience—his purposes, assumptions, and the effectiveness of his actions. "By assumption is meant that generally unconscious aspect of the transactional process which may be described as a weighted average of past experience in dealing with those portions of the impingements from the environment to which it is related. Assumptions function as probabilities which are built up by action, checked by action, and modified by action, as the consequences of these actions are registered in relation to purposes. Taken altogether, our assumptions form our 'assumptive world' which we bring to every occasion and on which our perceptions are based; therefore, the only world we know is determined by our assumptions" (p. 89). Or, in short, "what we see are constructs built up through experience" (p. 14). In the making of a 'choice' the individual may be in 'conflict' as to which one of the infinity of possible configurations is to be selected. In such cases the perceptual outcome apparently is to be explained in "terms of averaging our past experience" (p. 8). Here is a more specific statement which bears on the same point: "a perception is a prognostic directive for purposeful action. The reliability of this prognosis will be related to the reliability and number of the indications or abstracted aspects whose integration constitutes the perception. Presumably, the relative reliability of these indications is learned from experience while carrying out purposeful activity" (p. 185).

The foregoing theoretical statements underlie all the demonstrations and experiments. The monocular 'distorted' room may serve to exemplify the theory. Ittelson and Kilpatrick write: "Each part of a distorted room must subtend the same visual angle at the nodal point of an eye placed at the chosen observation point as would

the corresponding part of an undistorted room observed from the same point. If this prescription is followed, the stimulus-patterns on the retina will be the same for the distorted and undistorted rooms. In this fashion, a whole 'family' of distorted rooms of vastly different shapes could be built, yet all of them, if our reasoning is correct, should be seen as 'normal' instead of some other way because our past experience has made it a 'better bet' to perceive level floors, upright walls, rectangular windows, etc." (p. 48). Furthermore, it is assumed that social and personal factors enter into the process of perceiving the room and its contents (pp. 48, 239).

Some of the main results reported in the monograph which apparently confirm the underlying theory are as follows: (a) With regard to the 'distorted' room, all *Os*, under the conditions of the experiment, describe the room as "normal or very nearly normal" (p. 48). Moreover, it is suggested that as *O* explores the distorted room "he sees the room more and more in its true shape even though the stimulus-pattern on his retina has remained unchanged" (p. 51). (b) It is reported that in "size-distance" judgments of the projected images of "friendly" and "unfriendly" faces, the "friendly" faces are seen as "larger" than the other faces (p. 138). (c) "Relatively insecure" *Os* localize objects (pictures of a Rorschach inkblot, of a laborer, of an expensive automobile) as being nearer to them than do "secure" *Os* (p. 150). (d) Perceptual effects produced by a cylinder inserted at a given angle through the mullions of the rotating trapezoidal window depend on the assumption made about the nature of the cylinder. The *Os* reported a greater degree of distortion in the cylinder when they assumed it was of steel than when they assumed it was of rubber (p. 170). (They did not know that the same cylinder actually was used in both phases of the experiment.)

It may be well to raise here some questions which should be considered in a future reformulation and revision of the theory:

(1) Let us deal first with the main theoretical point: the fact that a given retinal pattern may be produced by an infinity of external configurations implies that the perceiver, on the basis of past experience, chooses and sees a particular one. This would mean that if an individual is reared in a world in which there are no 'rectangular' rooms or 'rectangular' forms he would see the 'distorted' room as actually distorted; furthermore such a hypothetical individual would not see oscillation in the rotating trapezoidal window. Yet a major empirical result is that *all* individuals (apparently, regardless of variations in past experience), under the given experimental conditions, see the 'distorted' room as normal and oscillation instead of rotation. The universality and uniformity of results suggest that the given perceptual process is a *necessary* one rather than one *chosen* by the perceiver. More attention should be given to the framing of a decisive experiment in this regard. For example, the "assumptive complex" or "weighted past experience" of *O* should be changed while the retinal pattern is held constant. Would there be any corresponding perceptual changes? This point, of course, is acknowledged by the Ames-Cantril theorists (p. 164). As mentioned earlier, it is suggested in the monograph that exploring the distorted room with a stick causes *O* to "see the room more and more in its true shape" (p. 51). The reviewer would venture the prediction that when the results of such an experiment are fully reported they will not fit the

Ames-Cantril theory. Parenthetically, the reviewer would like to see the phrase "true shape" clarified because another phase of the theory asserts that we can never know the actual or true character of the world.

(2) When *O* is properly placed in relation to the 'distorted' room he reports room which is 'normal' in appearance. An explanation in terms of past experience seems to be unnecessary since the 'distorted' room was specially constructed so as to produce the retinal pattern that a 'normal' room would produce (p. 48). Furthermore, suppose that the retinal pattern were not that of a 'normal' room. Would *O* then see a 'normal' room? The answer is obviously *no* because when *O* steps back from the viewing point in the distorted room demonstration he no longer sees the room as 'normal.' Shall we assume that there was a corresponding change in the "assumptive complex" of the observer? As the reader undoubtedly knows, the perceptual effects under consideration can be photographed. Does not the fact of photography imply that *O*'s perception is a necessary one? The observer's perception of a 'normal' room can be checked by actual measurements on the photograph; and, for this particular demonstration, measurements and perception coincide.

(3) When the cylinder is placed through the mullions of the rotating trapezoidal window distortion effects are a function of the "assumptive complex" of the observer. If the cylinder is assumed to be of steel (due to informal instructions of the experimenter) it distorts to a lesser degree than when it is assumed to be of rubber (p. 170). Accepting this experimental result (no statistical test of significance was reported), we may ask several questions. What would happen perceptually if *O* were led to assume that both the cylinder and the window are of steel? Why was it so easy to change the "assumptive complex" of *O* in this situation while it was so difficult (if not impossible) in other situations? Up to the time of the writing of the monograph the Ames-Cantril group has had ample experience in dealing with the 'true' nature of the objects used in the various demonstrations. They design, construct, constantly utilize and discuss the various pieces of apparatus. Has this close and prolonged experience with the experimental setups changed their own perceptions? Do they see the 'distorted' room in its true shape or the trapezoidal window rotate rather than oscillate? A reader of the monograph would gain the impression that, despite the knowledge of the members of the group, they observe the same phenomena as do their *O*s.

(4) There is sometimes insufficient clarification of experimental procedures and of the relation of a given experiment to the theory. For example, projected images of "friendly" faces are seen as "larger" than "unfriendly" faces (p. 138). How is such a result specifically related to the underlying theory? Moreover, this experimental result represents an inference of the experimenter on the basis of distance-judgments of *O*—"friendly" faces were set nearer to himself than were the "unfriendly" faces. Why weren't *O*s instructed to compare sizes rather than distances? In addition, the actual instructions used emphasized the subjective nature of estimates (p. 132). Would more 'objective' instructions produce different results?

(5) It is asserted that "man can never know more of the external world than those aspects which are directly relevant to the carrying out of his purposes" (p. 88). How is it that man apparently always has been able to see stars and stellar constellations when such objects seem to lie outside the scope of purpose and of action? Furthermore, how would the proponents of the Ames-Cantril theory ex-

plain those consistently reported visual phenomena about which the observer does not know, and is in no position to know, what the effects ought to be (for example, figural after-effects)?

(6) Does the transactional approach apply to lower animals? If some lower animals have perceptual responses which are basically similar to those of humans, are we to impute the same "assumptive complexes" to them?

(7) The reviewer notices no discussion of, or reference to, Pratt's (*J. Psychol.*, 30, 1950, 85) alternative explanation in terms of judgment of Hastorf's experiment (*J. Psychol.*, 29, 1950, 195) which apparently demonstrated an influence of meaning on distance-perception. The Hastorf paper is, however, included in the monograph. In a systematic elaboration of theory there should be some discussion of the Pratt's interpretation regardless of whether it is accepted or rejected.

The authors of the monograph are to be commended for bringing together their main results and for dealing with the theoretical issues generated by their results. The reviewer raised questions which he thought would occur to any careful reader of the monograph and which should receive some answer in further development of the theory.

Queens College

NICHOLAS PASTORE

The Sensations: Their Functions, Processes and Mechanisms. By HENRI PIÉRON. Translated from French by M. H. PIRENNE and B. C. ABBOTT. New Haven, Yale University Press, 1952. Pp. xxiii, 469.

The Human Senses. By FRANK A. GELDARD. New York, John Wiley & Sons, 1953. Pp. x, 365.

Piéron's book is an erudite personalized handbook of sense-physiology, written from a functional point of view and already somewhat dated.

Of Piéron's erudition no reader can remain long in doubt. The 793 items in his formal bibliography, the 700-odd names in his author-index, cite researches from 1790 to the present, 10% of them before 1900, 40% before 1930, most of them current. Quite often a crucial determination or a best measurement turns out to have been made long ago. Yet the selection is personal, Piéron's own. Any competent sense-physiologist or psychophysicist will be troubled by the omission of reference to many researches which he thinks are relevant. For instance, Geldard's book, with its much wider scope, lists 330 references against Piéron's 793 with only 44 common to the two lists. Only three of Geldard's citations are to French researches; 208 of Piéron's are French. Geldard cites 33 articles in German, Piéron 221. Geldard has 89% of his citations to English articles, Piéron 45%. Piéron has, moreover, many xenotypical statements, like his use of the phrase "Fechner's paradox" to indicate brightness-constancy instead of the failure of binocular summation, as we use the term in America. So the reader must remember that he is listening to Piéron, professor in the Collège de France, lecturing on what he finds important in his own very special field.

Within that special field this book is a handbook. It is replete with data, especially with times. Never before did the first thousand milliseconds of a sensation's life get such detailed analysis and description. You can read a chapter and gain from it a certain perspective of a limited set of events; yet there are in it so many nu-

merical facts that inevitably you come away with the major impression that here is the place to look a fact up, a set of standard tables on sense-excitation to be used for reference, not held in memory.

Nor does this text deal with experimental method or, except in the matter of the "Bouguer-Weber Law," with controversy. At last psychologists have a factual book to read, one that discusses the phenomena of sensory excitation, not the researches by which the phenomena were found. Piéron will indeed tell you that an exact value is not yet sure, that someone else got a different figure, that Fletcher's value for the threshold of hearing is ten times Wien's, but then he takes the accepted value and runs it through to all the consequences, imbuing the discussion with the spirit of certainty rather than the atmosphere of operation-dependent dubiety to which American psychologists have become accustomed.

Piéron's field is sense-physiology. It is scarcely psychology at all. The excitatory events at the receptors concern him most and it is only in the last chapter that he ventures so far afield as the brain. There the neurones that have to do with sensation are called "esthesiones" and what goes on consists of the recruitment of esthesiones by the excitation of numbers of fibers and by the frequencies of excitation in these fibers. The more esthesiones recruited, the more intense the sensation. You feel, however, that the careful physicalism, which Piéron finds adequate for getting from stimulation through the receptor and onto the excited fiber, halts confused before the psychocerebral mystery of an implicit dualism. Evidently Piéron belongs to the older generation that accepts, as the ultimate explanation of the psychophysical event, the correlation of an excited spot in the brain with the occurrence of a sensory experience. The newer generation notes, on the contrary, that no spot in the complicated excitatory course is less necessary than another, that a sensation as a scientific fact has to become public information, that its publication is an efferent event, and that in this manner physicalism becomes adequate to the explanation of sensory quality; but of this kind of consideration Piéron says nothing. Perhaps for him it smacks too much of American behavioristics.

The general intent of this book is to present a functional view of sensation. Piéron speaks constantly of excitation as messages, and actually the philosophy of information theory could have been used here with good effect, partially, indeed, to solve John Locke's as well as Piéron's problem of how the mind gets to know about the external world. The original title of this book was *Aux sources de la connaissance; la sensation guide de vie*. That title on its paper cover suggests a popular book for the multitudes, but Piéron is tough reading and the Yale Press did well to change the title and put on a cloth cover to warn Americans. Unfortunately Piéron's book is somewhat dated in this active field. The author wrote his first preface in 1944; the French edition came out in 1945; then Piéron made slight revisions for the translation which, with a preface dated 1950, was published in 1952. Some of the items which the sophisticated finds missing are researches that appeared too late, after 1944.

Piéron is clear-headed and scientifically tough throughout his exposition, except for his historically respectable softness in facing the problem of the scientific nature of sensory experience and—or so it seems to this reviewer—in one other matter. Piéron knows how difficult it is to state the nature of the intensitive stimulus for all the senses. For vision and hearing the stimulus can be expressed in terms of

energy, as Piéron says, or better in terms of change of energy, but the case is at present not such for smell, taste, touch, warmth, cold, pain and kinesthesia. He fights off all the soft-minded theorists who argue for the logarithmic relation between stimulus and sensation in respect of the Fechner law, even after they themselves have revealed exceptions to it and deviations from it, but then he concludes unequivocally for the existence of logarithmic relation in the receptor: "The intensity of excitation at the level of the receptor mechanism is proportional to the logarithm of the stimulus intensity" (p. 381). How could any such statement be possible without precise specification of the unit of measurement of the stimulus? And, if the relation is asserted for any one measure of the stimulus, it is *ipso facto* denied for every other measure not linearly related to the first. It seems to this reviewer that Piéron has here fallen into the same trap as his predecessors—an easy mistake, it is true, since the gross fact of diminishing returns for increasing intensity is so obvious for many measures of the stimulus although the precise logarithmic relation is not.

Geldard's book is a conventional handbook in the field of sensation, written with beautiful clarity at the level of the American graduate student in experimental psychology. This author's citations are more up-to-date than Piéron's and he also seems to need to rely less often on ancient researches. Only 3% of his 330 references are dated before 1900, 20% before 1930 as against Piéron's 40%, and 40% after 1940 as against Piéron's 20%. Geldard covers the field of sensation defined by Wundt as applying only to the facts and laws of sensory quality and intensity, with the spatial and temporal aspects of sensory experience left to be considered under perception—a dichotomy arising out of Wundt's respect for Kant's categories. Here, then, lies the major criticism of this volume: it leaves out all of space perception—visual, tactual and auditory, thus everything about form and the extensive organization of perception, all of temporal perception, thus the perception of time, the times of perception, the temporal organization of experience. Even Piéron has twenty-one pages on "the appreciations of time." The modern American view in textbooks and handbooks has come to be not to distinguish between sensation and perception, and the pity is that Geldard could not have written twice as big a book to include the entire field, thus stimulating the revival of courses in this oldest, most complete, and most scientific field of psychological instruction which is now losing student-interest. Unfortunately the present text is not big enough for a full semester-course.

Geldard's volume is more of a text than a handbook. Its five chapters on somesthetic sensation, a little more than a third of the book, are the fullest in relation to the state of knowledge, and also they are the best. Here lies Geldard's *Fach*, and we can be grateful that he has not restrained himself in this usually neglected field, even if his enthusiasm does unbalance the exposition a little over against vision and hearing, which together have somewhat less than half the total available space. Taste and smell get the remaining third, a very full treatment for each, but one that deals mostly in unresolved problems. Geldard could have afforded to cut here; yet this reviewer would prefer no cutting at all, just expansion for vision and hearing and then the addition of all of the space-time problems and facts.

Geldard discusses method, controversy, and unresolved issues as Piéron does not; yet he succeeds in impressing the reader with the healthy state of research in

three of the five sense departments. Some of his success in presenting dubious fact with interest lies in his skill in the use of the English language. He is by now one of the best writers among American psychologists. Piéron in English translation is tough and remains an enemy for the reader to conquer. Geldard is the reader's friend, the perfect teacher whose technique is so good that you never think about it until you start to write a review of him.

It is strange that two books with such similar titles could have so little content in common, but Piéron is writing sense-physiology and sticks to the periphery, whereas Geldard is writing "outer psychophysics," presenting correlation of sensory experience with stimulus-conditions, the experience that Piéron seems not to know how to capture in the net of science.

Have these books no errors? Little ones, and perhaps others that the reviewer is not wise enough to discover. Totaling the author's minor errors and the reviewer's criticisms, dissents, and regrets about omissions, we come out with less than three-score items for Geldard's 140,000 words, and about as many for Piéron's less familiar 190,000 words. The principal comment to be made about these books is that we are fortunate in having them both, fortunate in having this definitive account of Piéron's distinctive and distinguished contribution to science put over into English so that Americans can read it, fortunate in having Geldard making his wisdom, his maturity, his insight, and his skill in English composition available for the education of American psychologists, so many of whom are already forgetting that there is no topic in psychology in which there is so much accumulated knowledge and certain fact as the topic of sensation.

Harvard University

EDWIN G. BORING

The Conceptual Framework of Psychology. By EGON BRUNSWIK. International Encyclopedia of Unified Science, Volumes I and II: Foundations of the Unity of Science, Volume I, Number 10. Chicago, The University of Chicago Press, 1952. Pp. iv, 102.

This monograph comprises a broad survey of psychological problem-areas and methodology. A decrease in Brunswik's customary condensation makes for an increase in ease of reading, as well as for some sacrifice of the conceptual detail found in his earlier writing, to which, however, convenient references are made. The increasing popularity of the general viewpoint which Brunswik has explored so carefully makes this survey an especially important one today and, as a review of psychology's problems, it is imaginative and intriguing.

Brunswik approaches the problem of achieving 'objectivity' through the historical attempts to achieve epistemological *certainly*. With the Gestaltists, he rejects the constancy hypothesis (of a "one-to-one correspondence of sensory stimulus elements and conscious sensations," p. 4) but, unlike the Gestaltists, he makes no appeal to *organization*, and considers the relation of stimulation and experience to be an ambiguous one. With ambiguity therefore inherent in all our knowledge, "objectivity" is statistical, depending on the degree to which "... a class of responses yielding maximum reliability coefficients within or between individuals facing a ... common ... situation" (p. 11) is achieved; judgments as to the coincidence of points in space-time are the most reliable, and are therefore to be the language of

science. The correct level of behavior to be studied by "objective psychology" shows "vicarious functioning," *i.e.* the achievement of the same goal through different motor acts (detour behavior and so forth) and the perception of the same distal object even though the proximal sensory stimulus varies (the perceptual 'constancies'). The reliability of achieving a goal through any one motor sequence is low, as is that of perceiving distances of objects correctly through any one "depth cue," but objects are normally quite well perceived and attained. Nevertheless, since perceptual and behavioral achievement is imperfect, "the traditional nomothetic search for strict laws becomes an insoluble task" (p. 28), and psychology should study molar probability relations of a "central-distal" nature, assuming that "research should focus wherever behavior . . . may focus" (p. 25). We should eschew the admittedly more precise information concerning the simpler intervening (merely "mediating") processes with which psychologists have too frequently concerned themselves. To permit probabilistic vicarious functioning to manifest itself fully, we must achieve *situational representativeness*, since "the study of functional organism-environment relationships . . . require(s) that . . . focal events and other situational circumstances should be made to represent, by sampling or related devices, the . . . conditions under which the organism has to function" (p. 30). Thus, *representative* experimental design (analogous to representative random test sampling from a defined population of individuals) is to replace traditional *systematic* experimental design; generality of findings is to be increased by widening the samples of situations investigated.

Surveying current theoretical approaches, Brunswik finds the above points emerging in various areas of psychology (although some schools have become "encapsulated" in "microscopic" or "mediational" problems). His analysis is cogent and thought-provoking, despite a tendency toward oversimplification of schools to fit his classificatory schema, but this reader feels some misgiving concerning how much of this survey is relatively atheoretical *methodology* and how much is substantive (if implicit) theory. The answers a science gets depends upon the questions asked, and any dicta as to *what* and *how* to ask deserve careful scrutiny.

(1) "Objectivity" was defined as inter- and intra-individual observational reliability, and only "point-coincidences" and the "language of things" are presumably reliable enough. Should this not be an empirical question rather than an *a priori* decision? Other qualities of experience may yet prove no more variable. Brunswik agrees that a definition of "objectivity" need not reject introspection and (emphasizing the "basic deficiencies" of introspection in general by the difficulties of Würzburg phenomenology with the problem of imageless thought) shows us that thought *can* be communicably described—albeit only by language of "artistic character" (p. 5); yet he considers all introspective data as non-communicable and "private" to the "helpless solitude of the observer" (p. 4).

(2) Brunswik tells us that attention to *probabilistic* "molar" central-distal relationships will insure relevance and generality of conclusions; this would seem to require two assumptions. The first is that there is no one-to-one correspondence between experience and the proximal stimuli as subjected to *any* system of analysis; but ambiguity is not so all-encompassing, and the search for units of stimulus-analysis which *do* correspond with experience is by no means abandoned (*cf.* Gibson, Stevens). Brunswik rejects the traditional constancy hypothesis, but does not

offer any replacements for the traditional analytic units with which it was intertwined; not surprisingly, some ambiguity remains, to which he grants clear title to all Psychology. The second assumption would seem to be that we cannot understand "molar" phenomena by way of the part-processes; this is a not-unfamiliar position, but, again, not an atheoretical, neutrally representative one.

(3) *Representative* experimental design (i.e. "ecological sampling" of stimulus objects) will assure testing normal (and hence most generalizable) adjustive functioning; the logic is presumably the same when applied to generalization about populations of stimulus-situations as when applied to generalization about populations of organisms. Brunswik feels that he is dealing with a general methodological principle, citing Hammond (*Phil. Sci.*, 18, 1951, 211) who considers Brunswik's "theory of experimental design . . . to be congruent with Einstein's Theory of Relativity and 20th-century physics. Brunswik . . . and Einstein . . . appear to emphasize the same point, e.g., definition of the conditions of application of results is equally as important as definition of the conditions of achievement of results."

Is it here implied that with *representative* experimental design physicists would have known earlier that Newton's laws do not fit? Actually, of course, neither Einstein's nor Newton's laws are particularly derivable from ecologically representative situations. Physics, as all science, is abstractive, seeking special observational conditions; the corroborative Einsteinian data are drawn from such "unrepresentative" stimuli as discrepancies in Mercury's orbit. Brunswik probably has a particular *psychological* theory which makes such sampling procedures reasonable, and which can stand on its own feet; for example, representative sampling would improve the generality of data if organisms mirrored their environments, if whatever had a high probability of being associated in nature had a high probability of being "associated in" the psychological processes of the organism (an hypothesis with considerable historical precedent, surely explicit in the probabilistic empiricism of Hume and in Hartley's frequency-based associationism, but *not* a neutral, representative methodological principle).

In short, this is an intriguing overview of psychological concepts and methodology from a strong, consistent (and hence necessarily selective and biased) theoretical position, and will serve as an excellent introduction to Brunswik's viewpoint for those with an adequate psychological background.

Cornell University

JULIAN HOCHBERG

Psychoanalysis and Personality. By JOSEPH NUTTIN. Translated by GEORGE LAMB. New York, Sheed and Ward, 1953. Pp. xiv, 310.

The author, Father Nuttin, is a professor at the Catholic University of Louvain. His orientation is, however, in no sense exclusively European, and he is thoroughly familiar with virtually all of the relevant contemporary American literature, e.g. with the works of J. G. Miller, N. E. Miller, Dollard, Rogers, Snyder, Allport, Sears, Sheldon, Kinsey, and so forth. The book is divided into two major parts: a survey and criticism of orthodox Freudian psychoanalysis, and a presentation of the author's own 'dynamic' theory of personality. It also contains an outline of Adler's individual psychology, and an annotated bibliography of more than 125 titles.

Professor Nuttin's criticisms of Freudian theory are largely of the sort already

familiar to most psychologists. Thus Nuttin, like the neo-Freudians, argues against Freud's pan-sexualism (especially as illustrated by his extreme applications of the concept of sublimation), his excessive stress on infantile experiences as determining adult character and behavior, his over-emphasis on instinctive as contrasted with cultural factors, and so on. The author also contends that much of modern psychology is oriented far too strongly toward the abnormal; that it is too prone to interpret the statistically normal in terms of the clinical—witness Freud's identification of religion with an obsessional neurosis "mitigated by being so common" (p. 170); that it is too apt to expand concepts derived from clinical investigation into a general philosophy of human life.

Father Nuttin's own theory of personality is of the dynamic type in that it lays as much stress as possible upon motivational factors, at least as much as is consistent with the "rehabilitation of the cognitive functions in the psychology of human needs" (p. 161). It is, however, avowedly a theory meant to satisfy the "spiritually minded psychologist," the psychologist who "is convinced that the so-called 'higher' or *specifically human* elements in the psychic activity of man are not to be *reduced* to the 'lower' levels of behavior" (p. 38). To this end Nuttin makes throughout a sharp division between animal and human, between motivation through merely biological drives and through such dynamisms as "the drive to self-realization" and "a drive toward a certain self-transcendence to go beyond the merely spontaneous development of the biological process of growth" (p. 172).

In spite of his objections to many of Freud's conclusions, Father Nuttin explicitly states that "it is not by rejecting psychoanalytic doctrine but by making up for its deficiencies in a positive way" that we may create a fuller picture of human personality (p. 254). His aim is largely to effect a synthesis of the much that is left in psychoanalysis after his excisions and the convictions of a psychologist to whom moral and spiritual values have the status of imperatives. In trying to achieve this union, he probably has succeeded as well as the inherent difficulties of the task permit. His work is free from dogmatism, and arguments are presented with fairness as well as with much expository skill. The book can be recommended highly to all those who share the author's necessary premises.

New York University

LELAND W. CRAFTS

Hypnotism. By ANDRE M. WEITZENHOFER. New York, John Wiley and Sons, 1953. Pp. xvi, 380.

This book is intended to separate myth from actuality and fact from unfounded belief in the field of hypnotism. In the process Weitzenhoffer makes reference to over 500 studies. The text consists of three main parts: orientation, experimental foundations, and theory. Each chapter in the first two parts includes a survey of research together with an evaluation by the author and a concise summary.

On the positive side Weitzenhoffer makes the following important contributions: a definition of terms, a description of the different kinds of suggestion, a clear statement of the fact that present-day tests of hypnotic depth assume to begin with that *S* is hypnotized, and the separation of data obtained from experiments with hypnosis *per se* from experiments which make use of suggestion in hypnosis. Many of the older studies are reported but taken with a "grain of salt," as the author

points out, because of faulty experimental designs. The clinician will find many valuable leads in the book, *e.g.* the discussion of susceptibility to hypnosis of normal, neurotic, and psychotic individuals. Weitzenhoffer points out that most theories have assumed a simple cause-and-effect relationship. He proposes a multiple-factor theory whose key concepts are ideomotor activity, homoactive and heteroactive suggestibility, dissociation of awareness, and abstract conditioning.

Criticism tends to be of minor points. In certain areas interpretation of fact and fact itself are sometimes open to question. One may also be concerned that the theory proposed accounts better for experimental than for clinical data although the latter are by no means excluded. The author, although distinguishing among different types of regression, is not always clear as to whether he is using his own concept or that of the study under scrutiny. At times the ramifications of the author's theory and the maze of reduction sentences are so confusing that one is not always certain whether reader or author has lost the trend of argument. In sum, Mr. Weitzenhoffer is to be congratulated on writing a book which deals so adequately with the involved experimental and theoretical facts of hypnosis. He has sharply delineated many problems which previously had been ambiguous and obscure. It is the reviewer's opinion that the book is well worth the reading and can profitably be used as a graduate text in the field of hypnosis.

State College of Washington

F. L. MARCUSE

Murder, Madness and the Law. By LOUIS H. COHEN. Cleveland and New York, World Publishing Co., 1952. Pp. 173.

The author of this book is a psychiatrist of long standing who has taught psychiatry in the medical schools of the University of Illinois and Yale University, has been a clinical director in various State Hospitals, and at present is in private practice specializing in psychiatric problems of law. The purpose of the book is to show the differences in the points of view of jurist and psychiatrist in dealing with the psychotic individual who has committed murder. The author tries to show how legal standards can be set up that will resolve these differences and lead to more humane and just treatment of the offender.

The method is to recount a number of actual cases, to describe the behavior of the suspected psychotic person both prior to and subsequent to the performance of the murder, and then to follow the case through the courts. In this way the author deals with schizophrenia, malingering, temporary insanity, extreme frustration, and so forth. He does not approve of the concept of temporary insanity, and he does not think that a person who fakes insanity can mislead a competent clinical psychiatrist.

Not all psychiatrists are good witnesses; the best is the clinical psychologist of wide and continuous experience in diagnosis. The administrator who has lost contact with individual cases, the psychoanalyst who works with the neurotic rather than the psychotic, and the general psychologist do not make good witnesses. The clinical psychologist may contribute by interpreting tests and measurements of the defendant's personality and intellectual functions. The social worker and, where there is evidence of physiological disorder, the neurologist may also be of assistance.

The author writes clearly, logically, and without rancor against the courts. The

legal tests for insanity are bad, and until they are improved the best solution is to settle the issue of insanity before the psychotic is brought to trial. With good will, the differences between the courts and the psychiatrist may be solved. Everyone who is interested in its problems can read the book with profit.

Cornell University

H. P. WELD

Design for a Brain. By W. ROSS ASHBY. New York, John Wiley and Sons, 1952. Pp. ix; 260.

In this work the author seeks "to deduce from the observed facts what sort of mechanism it must be . . . that produces adaptive behavior." Although empirical data are alluded to occasionally, the method is primarily postulational.

The key conception about which the system revolves is the 'principle of ultra-stability,' which, stated formally, reads: "an ultrastable system acts selectively towards fields of the main variables, rejecting those that lead the representative point to a critical state but retaining those that do not." Behind this notion lies the simpler idea of a 'stable field.' A stable field exists when a certain variable under scrutiny (e.g. temperature under the control of a thermostat) stays within limits set by the designer. If, in the case of a stable system, the variable is forced outside the essential limits the stability of the system often breaks down (e.g. Weiner's example of the run-away steel mill). Under such conditions an ultrastable system will undergo a change of field, i.e. certain parameters of the system will vary until a stable field is generated. The crucial facts about these changes in parameter are that they only occur when the main variables go beyond essential limits and that the changes are in the form of step-functions. Ashby uses for illustrating this property a device of his own design (the Homeostat).

While the author claims to seek the sort of mechanism that *must* underlie adaptive behavior, like other *black-box* theorists he seems more concerned with the sufficiency of his theory than with its necessity. The unintelligibility of the definition of 'ultrastability' given above illustrates one of the principal difficulties of the book. The reader is forced to learn a new language before he can follow the arguments. While the development of such a new language is often a prerequisite to scientific advance, it does not appear to have yielded sufficient clarity to be justified in this case.

University of Texas

JOHN KRAUSKOPF

A Manual for Psychiatric Case Study. By KARL A. MENNINGER. New York, Grune and Stratton, 1952. Pp. xii, 355.

There is no doubt that one of the major weaknesses in the preparation of clinical psychologists has been in the area of case-reports. One has only to listen to a presentation of case-material before a staff conference or read the results of a psychological examination, to realize the inability of some psychologists to present a clear, concise, and meaningful description of a particular patient. Though Dr. Menninger's manual has been written primarily for the doctor entering the field of psychiatry, many of his criticisms apply equally well to the clinical psychologist. Particularly pertinent are the author's comments in Part III on the writing of psychiatric reports and papers: "If he is to write clearly and accurately, then, the psy-

chiatrist must be aware of the inferential and hypothetical nature of many psychiatric concepts and must distinguish clearly between his observations and the inferences he draws from them . . . verbosity and technical jargon in psychiatric reports are frequently symptoms of unclear thinking or of insufficient attempts to compensate for his own lack of knowledge by indiscriminately flooding the page with words, in the hope that a few of them will hit the mark" (p. 163).

Over half the book is devoted to actual case-reports illustrating the procedure for collection of information and writing of the report. The material is interesting, but one report might have served as well as four. The objective of the case-history, as the author states, is to indicate and justify a specific therapeutic program, yet nowhere do we see the connection between many of the data collected and the recommended therapeutic programs. One can become all too preoccupied with 'data-collecting!' A useful addition to the text is the appendix on the psychiatric nomenclature of the American Psychiatric Association and the Veterans Administration, a list of personality types, and adjunctive therapy modalities. A copy of a model commitment-law, drawn up by the Federal Security Agency of the United States Public Health Service, also is provided.

University of Texas

PHILIP WORCHEL

Basic Psychiatry. By EDWARD A. STRECKER. New York, Random House, 1952. Pp. 473.

This is a popular presentation of psychiatry by a well known personality in the field. One theme of the book is that mental patients "are much the same as we are" (p. 6). Another is that "Emotional maturity is the most important thing in the world" (p. 7). The first theme is developed by devoting chapters to the major psychotic groups, to psychoneuroses, psychosomatics, pathological drinking, mental defects and C. P. I. The problem of emotional maturity is treated in a chapter which indicts many contemporary mothers and sets the stage for a second chapter on "A design for childhood." The book resembles a popularized version of a text in abnormal psychology except that the word *psychology* is not found in the index. Neither are the phrases *social worker*, *projective test*, or *intelligence test*. The book varies little (except possibly for a chapter on the psychiatry of warfare) from one that might have been written 15 years ago. There is no bibliography. There is too little recognition given to the fields of science and service that border on psychiatry or the modern team-approach to cases. This defect is specially noticeable in the treatment of subjects, such as morale and intelligence, which lie outside the major province of psychiatry.

Many interesting brief cases and humorous anecdotes drawn from Dr. Strecker's wide experience will amuse as well as instruct the layman. The treatment of causation in mental disorders and the chapter on "functional sickness" ought to be particularly provocative.

University of Missouri

FRED MCKINNEY

Readings in Industrial and Business Psychology. By HARRY W. KARN and B. VON HALLER GILMER. New York, McGraw-Hill, 1952. Pp. ix, 476.

There has long been a need for a book of this kind. After a survey of "approximately 500 articles in the literature," the authors selected 53, because of "sound

writing and concrete contribution," for inclusion in this book. The articles are classified in eleven groups: (1) Motivation and morale, (2) Training in industry, (3) Analysis and evaluation of job performance, (4) Psychological tests, (5) Interviewing and counseling, (6) Accidents and safety, (7) Fatigue and worker efficiency, (8) Market research, (9) Industrial leadership, (10) Industrial relations, and (11) Psychologists in industry. The articles range from the purely arm-chair type to factual reports of carefully controlled studies. Each is directly reprinted and introduced with a four- or five-line summary by the editors. Although 53 other equally good articles could have been chosen, those selected are indeed representative, which is all that the editors claim. It might have been possible, however, by judicious condensation, to double the number of articles in the space allotted. Nevertheless, this book of readings should prove to be very useful, especially on the undergraduate level.

University of Georgia

J. STANLEY GRAY

Psychoanalysis as Science. By E. PUMPIAN-MINDLIN, ERNEST R. HILGARD, and LAWRENCE S. KUBIE. Stanford University Press, 1952. Pp. x, 174.

The Hixon Lectures on the scientific status of psychoanalysis delivered at the California Institute of Technology in 1950 are reproduced in this book. Kubie centers his discussion on the way in which the psychoanalytic *interpretation* is presumed to function in producing behavior change, pointing to the obstacles that must be overcome if this intervention on the part of the therapist is to be understood in a scientifically acceptable way. Hilgard examines the validity of certain psychoanalytic concepts by summarizing and evaluating experiments that have been conducted in order to provide an impartial test of them. Pumpian-Mindlin considers technical difficulties facing psychoanalysis in contrast with those facing the biological and social scientists in developing a science of behavior. The product is a well-written, nonpartisan book. It touches on many of the key problems that must be satisfactorily dealt with before psychoanalytic concepts can become a body of systematically verified propositions concerning human motivation and the conditions that produce predictable changes in behavior.

University of Texas

ROBERT R. BLAKE

Manic Depressive Psychosis. By LEOPOLD BELLAK. New York, Grune & Stratton, 1952. Pp. xii, 472.

Here is a systematic summary of the literature (about 1200 references) on the manic-depressive syndrome and related disorders, a companion volume to the author's earlier work on schizophrenia. Bellak's theoretical approach to the two psychoses is very much the same, but the volume at hand is presented only as a review of the literature and the author usually offers no judgment concerning the validity of the materials included. In the preface Bellak suggests the need for a central agency which would help to administer and to coordinate statistical research on psychiatric problems. He also offers the opinion that much work and much journal-space are wasted as a result of the physician's lack of experimental sophistication.

College of Medicine
Baylor University

FLOY JACK MOORE

The Story of the Adaptation Syndrome. By HANS SELYE. Montreal, Acta, 1952. Pp. 225.

Second Annual Report on Stress. By HANS SELYE and ALEXANDER HORAVA. Montreal, Acta, 1952. Pp. viii, 526.

The first volume, based on a series of recorded lectures, provides a refreshingly informal account of the inception and course of the author's researches on the "pharmacology of dirt." Psychologists who share the current interest in problems of stress will find here a valuable summary of Selye's ideas which can be followed with a minimum of physiological sophistication. The student will profit from the glimpses of science *in situ* which the volume affords. The *Annual Report* serves as a guide to more than 4,000 recent papers.

M. E. B.

Essentials in Interviewing. By ANNE F. FENLASON. New York, Harper, 1952. Pp. xi, 352.

This posthumously printed work is an attempt to bring together a few copiously illustrated principles of interviewing. The text was designed primarily for students of case-work and for others interested in clinical interviewing, and the material should prove interesting reading for the undergraduate student in these areas. We can only agree, however, with Miss Fenlason's own evaluation of the originality of her contribution. She quotes Montaigne in saying that "I have here only made a nosegay of culled flowers, and have brought nothing of my own but the thread that ties them together." For most readers familiar with the literature and techniques of interviewing, the treatment is likely to prove both uncritical and elementary.

Cornell University

PATRICIA C. SMITH

The Psychology of Learning. By E. R. GUTHRIE. Revised edition. New York, Harper & Brothers, 1952. Pp. x, 310.

To his well known treatment of learning, Guthrie has added four chapters—one on the "pluralistic theories" of Tolman and Maier, a second on reinforcement theory (Hull, Spence, Miller and Dollard), a third on Skinner, and a fourth on the study with Horton of cats in a puzzle-box. The relation between Guthrie's formulation and the great bulk of experimental work on learning remains remote.

M. E. B.

THE AMERICAN JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXVII

SEPTEMBER, 1954

No. 3

A VALIDATION STUDY OF RATIO-JUDGMENT METHODS

By J. P. GUILFORD and HARVEY F. DINGMAN,
University of Southern California

In recent years we have seen an increasing utilization of the fractionation method, and its reverse companion the multiple-stimulus method, in psychological-scaling studies. In the former, the observer, *O*, finds what he believes to be a stimulus giving a psychological impression of a defined nature that is some fraction of that given by a standard stimulus, the fraction usually being $\frac{1}{2}$, but also sometimes $\frac{1}{3}$, $\frac{1}{5}$, and so on, down to $\frac{1}{100}$. In the multiple-stimulus method the ratios are reciprocals of such fractions; *O* finds a stimulus that gives an impression *X* times as great as that given by the standard stimulus.

In both methods the assumption is made that *O* can report faithfully on ratios of stimulus-effects and that the values on the psychological scale resulting from such judgments are ratio-scale values, as defined by Stevens.¹ This assumption has been tested to some extent by experiments on the internal consistency of results from the fractionation and multiple-stimulus methods.² The findings have been generally reassuring, but in both of these methods the judgments have the common feature of *O*'s looking for a stimulus to fulfill a specified fraction or multiple. The test is therefore not as general as it might be. Using results from such sources can still be questioned. Since psychological conclusions of considerable importance are based on such results, we would do well to make further examination of the methodology.

There is another psychophysical method that utilizes ratio-judgments

* Accepted for publication August 10, 1953.

¹ S. S. Stevens, On the theory of scales of measurement, *Science*, 103, 1946, 677-680.

² For example, R. M. Hanes, A scale of subjective brightness. *J. Exper. Psychol.*, 39, 1949, 438-452; The construction of brightness scales from fractionation-data: A validation, *idem.*, 719-728.

in a different manner. This is Metfessel's constant-sum method.³ In this method *O* observes two or more stimuli in conjunction and essentially makes a report concerning the ratio that he believes exists psychologically between them. If stimuli *A* and *B* are presented and *O* divides 100 points in the proportions 80 and 20, respectively, he is telling us that to him the impression of *A* is four times that of *B*, also that the impression of *B* is $\frac{1}{4}$ that of *A*.

Our first interest in this study was to see whether the constant-sum method leads to the same scaling of stimuli as the fractionation-method. We were also interested in discovering more facts concerning the performance of the constant-sum method, since there is almost no publication concerning its validity or utility. We have applied the constant-sum method in two ways, one in which *O* judges two stimuli at a time and one in which *O* judges five stimuli at a time. We have also applied the method at somewhat different stimulus-levels. We were also concerned with the type of psychophysical law for lifted-weights that would be indicated by the results from both the fractionation- and constant-sum methods; whether all ratio-judgment methods lead to the same psychophysical law and whether this law is in agreement with previous findings regardless of method.

EXPERIMENTAL PROCEDURES

Five experiments were performed, one in which the fractionation-method was used and four in which the constant-sum method was used. The conditions can be described briefly.

In four experiments 10 standard weights were used, their values being 40, 100, 150, 200, and 300 gm. in a 'small' set and 300, 550, 900, 1400, and 2000 gm. in a 'large' set. It was not possible to vary weights from 40 to 2000 gm. in containers of the same size. The containers were metal, cylindrical, ointment boxes. Those in the small set were 5.0 cm. in diameter and 1.7 cm. deep. Those in the large set were 10.5 cm. in diameter and 6.7 cm. deep. We were able to check on comparability of results from the weights of two sizes. The results seemed always to fall in line. Only weights within a set, large or small, were compared by *O*, thus avoiding the size-weight illusion, it would seem. It is our opinion, however, that this illusion is due for further investigation in lifted-weight experiments.⁴ It will be noticed that the heaviest weight in the small series was identical in weight with the lightest weight in the large series. This enabled us to bridge the gap between the two in treating the results.

Experiment 1. In Experiment 1, *O* was to select a weight in a comparison series

³ M. Metfessel, A proposal for quantitative reporting of comparative judgments, *J. Psychol.*, 24, 1947, 229-235.

⁴ Even in a single set of constant size, the average densities of stimuli vary and the general level of density within a set may be an unnatural one for *O*. The possible effect of density on judgments of weights has been generally overlooked.

that seemed to him to be one-half as heavy as each standard. Instead of providing six comparison weights for each standard, as has been common practice,⁶ we presented one long, continuous series of comparison weights for use with all five of the weights in each set. The small comparison weights were: 20, 24, 28, 34, 40, 50, 63, 78, 100, 110, 120, 139, 150, 165, 175, 185, and 200 gm. The large comparison weights were: 150, 165, 175, 185, 200, 210, 245, 300, 350, 410, 475, 550, 630, 700, 790, 900, 1000, 1125, 1250, 1400, and 1550 gm. Thus, both the standard sets and the comparison sets were very roughly in geometric series.

Each of 20 *Os* (graduate students in psychology) found a comparison weight that seemed to him to be half that of each standard weight. Half of the *Os* started with the lightest standard and half with the heaviest in each set and worked serially through the set. Half of the *Os* judged the large set first and half the small set. We therefore have 20 judgments of S_h (the half-stimulus) corresponding to each standard S_s . For a single value for the half-stimulus we obtained a geometric median of the 20 judgments with that standard. Since further computations with the data were to be in terms of logarithms of S_s and S_h , it was regarded as most appropriate to average $\log S$. Since in these small-sample distributions there was considerable irregularity, particularly in the form of lone extreme values, the median was preferred to the mean.

Experiment 2. In Experiment 2, the standard stimulus-weights were presented to *O* in all possible pairs within each set. For each pair *O* was instructed to divide 100 points between the two stimulus-weights. Each of 20 *Os* judged every pair once. Thus, in effect, 2000 points were divided between members of every pair, from which a ratio could be calculated. The inter-pair ratios were treated by procedures described by Comrey⁶ in order to arrive at a single set of ratio-scale values for the stimulus-weights within each set.⁷ The two sets of scale-values were converted to a common unit by using the fact that the stimulus-weight of 300 gm. appeared in both sets.

Experiment 3. In Experiment 3 all five stimulus-weights in each set of standards were presented simultaneously and *O* was to divide 100 points among the five. The total numbers of points assigned to the respective weights by 20 *Os* can be taken to represent ratio-values, which can be adjusted to any arbitrary unit by simple multiplication. The common stimulus-weight in the two sets was used, in this experiment also, as a means of deriving a common unit. The unit adopted in these experiments will be explained later.

Experiment 4. In Experiment 4, we used the same two sets of standard weights in connection with two parallel sets of half-weights corresponding to them as determined by Experiment 1. The half-weights directly estimated from the judgments of Experiment 1 were actually not used. The half-weights used were estimated from the linear function relating $\log S_h$ and S_s (Equation [1]). This relationship is

⁶ R. S. Harper and S. S. Stevens, A psychological scale of weight and a formula for its derivation, this JOURNAL, 61, 1948, 343-351.

⁶ A. L. Comrey, A proposed method for absolute scaling, *Psychometrika*, 15, 1950, 317-325.

⁷ For descriptions of computing procedures used in this article, see the senior author's *Psychometric Methods*, 2nd ed., 1954, Ch. 9.

shown graphically in Fig. 1. The values of the half-weights used may be seen in Table V. In Experiment 4, each half-weight, S_h' , so determined, was paired with its corresponding S_s for constant-sum judgments. O was to divide 100 points between the two stimulus-weights. The object was to see whether for each pair and for pairs in general the average value would be 2.0 for the ratio S_s/S_h' .

Experiment 5. In Experiment 5, only the half-weights (see Table V) were used. They were presented in all possible pairs, as the standards were in Experiment 2, for judgments by the constant-sum method. The object was to see whether the same inter-stimulus ratios would be obtained as in Experiment 2, and whether when scaled to the same unit, the psychological values would be one-half those found in Experiment 2. The results also provided one additional opportunity to examine the psychophysical relationship between psychological and physical weight values, and to see whether all the methods lead to the same relationship.

RESULTS

Experiment 1. The first objective in connection with the data from Experiment 1 was to relate the experimentally determined half-stimulus

TABLE I
TEN STANDARD STIMULUS WEIGHTS AND THE GEOMETRIC MEDIAN OF WEIGHTS
JUDGED TO EQUAL ONE-HALF OF EACH STANDARD

	Standard weight (S_s)	Median weight judged half as heavy as standard weight (S_h)
Small test	40	23.3
	100	51.6
	150	83.3
	200	119.9
	300	165.0
Large test	300	173.2
	550	324.0
	900	543.0
	1400	801.9
	2000	1134.0

values to the standard stimulus-values. These data are shown in Table I. Logarithms of both S_h and S_s were obtained. Fig. 1 shows the plot of S_h as a function of S_s on log-log paper. The regression is definitely linear and the data from the two sets fall very nicely into line. A least-square solution yielded the equation

$$\log S_h = 1.0105 \log S_s - .2699. \dots\dots\dots [1]$$

The goodness of fit of this linear relationship is indicated by a product-moment correlation of 0.999.

The second objective was to derive a functional relationship between a psychological-weight scale and the physical-weight scale. In a similar

experiment, Harper and Stevens⁸ found that the relationship called for the equation

$$\log V = 14.58 \log (1 + \log S) - 6.94, \dots\dots\dots [2]$$

where V = psychological weight in a unit called the 'veg,' and S = physical weight in grams. The veg-unit was arbitrarily defined as the psychological quantity that corresponds to 100 gm.

To test the applicability of the Harper-Stevens equation to our data,

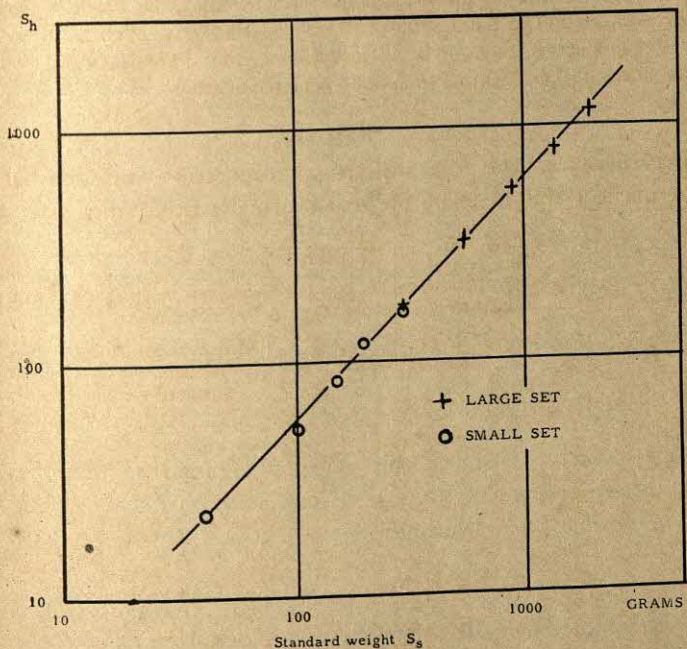


FIG. 1. RELATIONSHIP OF THE MEDIAN OF THE LOGARITHMS OF HALF-WEIGHTS S_h TO THE LOGARITHMS OF THE STANDARD WEIGHTS S_s , FOR BOTH LARGE AND SMALL SETS

it is necessary, first, to achieve the psychological scaling of the weights in vegs. We have as a starting point the fact that a stimulus-weight of 100 gm. corresponds to the psychological quantity of one veg. We also have the equation relating S_h to S_s . Working both ways from 100 gm., we found the stimulus-weights corresponding to 0.5, 0.25, and 0.125 vegs, in the one direction, and to 2, 4, 8, 16, 32, and 64 vegs, in the other direction. The results are shown in Table II.

⁸ *Op. cit.*, 346.

Having derived 10 pairs of veg- and gram-values covering the range of the experimental weights, we attempted to find the best-fitting equation of the Harper-Stevens type. A plot of $\log V$ against $\log (1 + \log S)$, however, showed definitely a nonlinear relationship. On the other hand, an essentially linear relationship was found when $\log V$ was plotted

TABLE II

DETERMINATION OF CORRESPONDING PSYCHOLOGICAL WEIGHTS IN VEGS AND PHYSICAL WEIGHTS IN GRAMS, USING THE RELATIONSHIP: $\log S_h = 1.0105 \log S_s - .2699$
(FROM EXPERIMENT 1)

V (vegs)	$\log S$	S (gm.)
.125	1.2454	17.6
.25	1.4996	31.6
.5	1.7511	56.4
1	2.0000	100.0
2	2.2463	176.3
4	2.4900	309.0
8	2.7312	538.5
16	2.9699	933.1
32	3.2061	1607
64	3.4399	2754

against $\log S$. By the method of successive differences we found the equation

$$\log V = 1.2346 \log S - 2.4692 \dots \dots \dots [3]$$

with the second constant in this equation adjusted so that the curve passes through the point (2, 0). Taking antilogarithms, we have the equation relating V to S , which reads

$$V = .00339 S^{1.2346} \dots \dots \dots [3a]$$

Experiment 2. Table III shows the psychological ratios of the weights in each set to the lightest weight in the set. Letting the stimulus-weight of 100 gm. correspond to the psychological unit of one veg, we have the five veg values for the small set given in the third column. Since the number of vegs corresponding to 300 gm. is 3.854, we multiply each ratio in the list for the large set by this constant to obtain corresponding vegs for the stimulus weights in that set. This is a somewhat risky procedure, since an experimental error in the scaling of the 300-gm. weight in the small set would be reflected in all the scale values in the large set. A check that was made on this, however, indicates no appreciable error. The scale values for the large set range from 3.854 to 33.847 vegs.

A plot of $\log V$ against $\log S$ again showed a linear regression. It also reveals no systematic discrepancy between small-set values and large-set

values, which is the check referred to above. A least-square solution based on nine pairs of observations (we lose one degree of freedom in using the ratio for 300 gm. to establish equal units in the two sets) gives the equation⁹

$$\log V = 1.1177 \log S - 2.1996. \dots\dots\dots [4]$$

Transformed into power form, the relationship of V to S is stated by the equation

$$V = .00632 S^{1.1177}. \dots\dots\dots [4a]$$

The index of correlation is 0.998. The fact that the same type of equation

TABLE III
STANDARD WEIGHTS, PSYCHOLOGICAL RATIOS, AND PSYCHOLOGICAL WEIGHTS IN
VEGS DERIVED FROM JUDGMENTS BY THE CONSTANT-SUM
METHOD IN EXPERIMENTS 2 AND 3

Stimulus weight	Experiment 2 (Judgment by twos)		Experiment 3 (Judgment by fives)	
	Ratio	Vegs	Ratio	Vegs
40	1.000	.452	1.000	.485
100	2.212	1.000	2.060	1.000
150	3.526	1.594	3.683	1.788
200	4.905	2.217	5.089	2.907
300	8.524	3.854	9.126	4.430
300	1.000	3.854	1.000	4.430
550	1.794	6.916	1.972	8.736
900	3.202	12.341	3.687	16.333
1400	5.457	21.031	6.335	28.063
2000	8.780	33.837	9.352	41.427

applies to the relation of V to S in both Experiments 1 and 2 is some indication that the fractionation and constant-sum methods achieve similar scaling results and the same general psychophysical law for lifted-weights.

Experiment 3. The essential results of Experiment 3 are given in the last two columns of Table III. The ratios within the two sets bear some resemblance to those derived from Experiment 2, but the range is somewhat greater. When transformed to the veg scale, the difference in range seems even greater. The two series of veg estimates bear a linear relationship when plotted, which indicates that the difference is merely one of

⁹ In this equation, and in those from Experiments 3 and 5, a minor adjustment would need to be made in the second constant in order to assure that the function passes through the points (2, 0) on the logarithmic scale, thus to assure the intended unit on the psychological scale. To meet this requirement, the second constant must equal (-2) times the first.

unit. The unit in either experiment depends upon one experimental value—the ratio-value determined for the 100-gm. weight. If there is an experimental error in this value, the unit is in error for the entire experiment. Linear transformation may be commonly needed in order to equate units for scale-values from different experiments.

The relationship of V to S indicated in the results of Experiment 3 is similar to those found in Experiments 1 and 2. A least-square solution based on nine pairs of observations gives the equation

$$\log V = 1.1760 \log S - 2.2695, \dots\dots\dots [5]$$

or in power form,

$$V = .00538 S^{1.1760}, \dots\dots\dots [5a]$$

The index of correlation is 0.998.

Experiment 4. Experiment 4 is a crucial one for the cross validation of the methods of fractionation and constant sum. It may be remembered that

TABLE IV

RESULTS OF JUDGING STANDARD WEIGHT S_s AND ITS MOST PROBABLE HALF-WEIGHT S_h IN PAIRS BY THE CONSTANT-SUM METHOD, IN EXPERIMENT 4

Standard weight (in gm) S_s	Mean of points assigned	t	Ratio of psychological correlates of S_s and S_h
40	34.80	0.77	1.87
100	33.05	0.13	2.03
150	34.40	0.61	1.91
200	33.05	0.16	2.02
300	30.15	1.99	2.32
300	36.15	1.40	1.77
550	34.75	0.61	1.88
900	33.30	0.01	2.00
1400	33.40	0.04	1.99
2000	30.55	1.24	2.27
All stimuli	33.36	0.01	2.00

in this experiment O was presented with 10 pairs of stimulus-weights to be judged by the constant-sum method. Each pair was composed of a standard weight S_s and its most probable half-weight S_h determined from the results of Experiment 1, in which the fractionation method had been used to derive the equation relating S_h to S_s .

The mean number of points assigned to S_h for each standard S_s is shown in the second column of Table IV. To satisfy the prediction, each mean should be 33.33. The means actually range from 30.15 to 36.15, with

a mean for all stimuli equal to 33.36. A t -test was made to see whether any of the means differed significantly from the expected mean of 33.33. The largest t , which was 1.99, is not significant at or beyond the 5-% level. We may therefore regard the hypothesis of agreement of ratios by the two methods as quite tenable for all stimuli.

In the last column of Table IV we see that the psychological ratios for the pairs varied from 1.77 to 2.32, with an overall mean of 2.00. That this range can be regarded as due to chance is indicated by the result of an F -test applied to the 10 means of column 2, Table IV. This test was applied after determining that intercorrelations of judgments of pairs over the population of individual O s were generally close to zero. The F -ratio was 0.91, where it requires an F of 1.98 for significance at the 5-% point.

Perhaps it is worth noting that the mean ratios for the small and large sets of stimuli were essentially the same, showing that the size of the weight containers, within the limits of those used, has no bearing upon the judgments of ratios of weights. Within each set, however, there seems to be a slight tendency for ratios for extreme stimulus-weights to be out of line, with the heaviest having the smallest ratios and the lightest having the largest ratios. Since the F -ratio was insignificant, too much should not be made of this apparent constant error, but it may be a lead for further investigation into the properties of ratio-judgments in the constant-sum method.

Experiment 5. Still another test of consistency of scaling results was to use the half-weights predicted from standard weights on the basis of Equation [1] from Experiment 1. The expectation was that the ratios among the half-weights would be very similar to those among the standards, and that the scale values in the same veg unit would be half as large as those for the standard weights. The half-weights were judged by the constant-sum method in every possible pair within each set. The ratios of all weights to the lightest in the set are given in the second column of Table V. The best comparison to make is that between these ratios and those obtained by the same procedure for the standard weights, in Experiment 2. The latter ratios will be found in Table III. For the small set, the ratios in Experiments 2 and 5 are remarkably similar, the greatest discrepancy being 0.019. For the large set, however, two of the ratios are in poor agreement, the largest discrepancy being 1.380. The reason for this is not clear.

It is also of interest to see whether a similar psychophysical equation will be found from the data in Experiment 5 as in other experiments.

Deriving ratios on a common scale for the two sets by using the one common stimulus of 171.1 gm., and treating these ratios as psychological values, by a least-square solution relating these values to S_h' we find the equation

$$\log R_c = 1.1121 \log S_h' - 1.5715. \dots\dots\dots [6]$$

The correlation was 0.996.

We want an equation relating vegs to stimulus-values. R_c is related to V by a constant ratio, or in the logarithmic equation by a constant incre-

TABLE V
THE MOST PROBABLE HALF WEIGHTS, AS DETERMINED BY EXPERIMENT 1, THEIR
PSYCHOLOGICAL RATIOS AND PSYCHOLOGICAL WEIGHTS IN VEGS

Stimulus (S_h')	Ratio within sets (R)	Ratio on common scale (R_c)	Vegs (V)	Ratio of vegs V_s/V_h'
22.3	1.000	1.000	.222	2.04
56.4	2.213	2.213	.492	2.03
85.1	3.528	3.528	.785	2.03
113.6	4.907	4.907	1.092	2.03
171.1	8.543	8.543	1.919	2.01
171.1	1.000	8.543	1.919	2.01
315.7	1.633	13.951	3.104	2.23
519.2	3.212	27.440	6.104	2.02
849.4	5.219	44.586	9.919	2.12
1164.0	10.160	86.797	19.309	1.75
			M	2.03

ment. Letting $\log S_h' = 2.0$ and solving the equation, we find that $\log R_c$ then equals 0.6527 and hence R_c then equals 4.495. Thus, an R_c of 4.495 is equivalent to one veg. Or, we can say that $R_c = 4.495 V$, or that $\log R_c = \log 4.495 + \log V$, or that $\log R_c = 0.6527 + \log V$. Substituting this last expression in Equation [6] and simplifying, we have

$$\log V = 1.1121 \log S - 2.2241. \dots\dots\dots [7]$$

In power form,

$$V = .00597 S^{1.1121} \dots\dots\dots [7a]$$

The conversion of the obtained ratios in column 3 of Table V to vegs was done and the results are given in the fourth column. The ratios of the veg values of the standard stimuli (Experiment 2, Table III, column 3) to the veg values of the half-stimuli column 4, Table V) are given in the last column of Table V. All except three are very close to 2.00 and the average of the nine independent ratios is 2.03.

DISCUSSION

While considering the outcomes of the five experiments together, it is well for us to keep in mind the major problems of the investigation. First, we wished to determine whether the method of constant-sum would yield results consistent with the method of fractionation. Both methods call for ratio-judgments on *O*'s part and treat the numerical data as if they are values on ratio-scales. The comparison of the two methods is of interest because in the former, *O* is presented with stimuli and he is to make a quantitative judgment of their ratios, while in the latter *O* is presented with a stimulus and a specified ratio that he is to fulfill by selecting an appropriate second stimulus. Secondly, we wished to determine whether different variations of the constant-sum method would yield consistent results with one another as well as with results from the fractionation-method. Thirdly, we wished to see what psychophysical law for lifted-weights would be found by means of ratio-judgment methods and whether the law found by Harper and Stevens would be verified.

Scale values from different procedures. Several kinds of tests were made in the attempt to solve the problems mentioned in the preceding paragraph. Except for Experiments 2 and 3, there was no opportunity to determine directly whether the psychological scaling of selected stimuli yielded concordant scale values. Table III shows ratios and scale-values that are similar but are by no means identical. One list of each kind of values (ratios and scale-values in veg units) was derived from judgments of pairs of stimulus-weights and the other from judgments of five at a time, both by the constant-sum method. Thus, from this limited finding, it appears that the method of constant-sum may yield variations in scale-values depending upon the number of stimuli presented simultaneously.

An indirect test of agreement between scale-values from different procedures is provided by Experiment 5, in comparison with Experiment 2. The veg values found for the half-weights are generally close to half those for the standard weights. In both experiments, however, *O* judged pairs of stimuli. Thus we may conclude that for sets of stimuli at somewhat different levels on the stimulus-scale, the psychological scaling will be concordant when pairs of stimuli are judged in all sets. How divergent the sets may be in stimulus-level we cannot say.

Psychophysical law from different procedures. We were able to apply more generally the test of whether the different procedures would lead to the same quantitative relationship between psychological and physical weight-values. The same type of equation was found from four of the

experiments, with similar parameters. It is worthy of note that the parameters for the three equations derived from the constant-sum method are very similar but all three differ somewhat from those derived from the fractionation-method. Whether this systematic difference will hold up when the fractionation-method is applied in other experiments we cannot say. The most similar pair of equations come from Experiments 2 and 5, in which the constant-sum method with paired observations was employed. Again, we need additional examples to see whether this is a function of method.

The psychophysical law for lifted-weights. The psychophysical law for lifted-weights as derived from ratio-judgments is clearly not the Fechner law. The exponent of S being greater than 1.0, the curve even has *positive* acceleration rather than negative as in the Fechnerian law. The Harper-Stevens function was not verified by these results, even when the same psychophysical method was used.

The power function obtained is what one would expect from the Guilford n th-power law. The latter was originally stated as a generalized substitute for Weber's law.¹⁰ The n th-power law in equation form reads

$$\Delta S = KS^n \dots\dots\dots [8]$$

where S = the stimulus-value, ΔS = some defined increment, and n = a constant that was expected to vary between 0.5 (the Fullerton-Cattell square-root law) and 1.0 (the Weber law). If we integrate Equation [8], as Fechner integrated Weber's equation, we obtain an equation of the form

$$R = CS^m + a \dots\dots\dots [9]$$

where R = quantitative value of a response, $m = 1 - n$, where n is the exponent in Equation [8], and a = the constant integration.

Equation [9] calls for a constant term a , which was not taken into account in the empirical equations fitted to the experimental data. In those equations we may say that it was assumed that a is zero. It is difficult to obtain an accurate estimate of this constant. In the functional relation of V to S we should expect a to be negative, for S should have some positive value when V is zero. Such is not the case in these lifted-weight data. Plots of V as a function of S show that a must be a positive quantity. In the various Experiments, 1, 2, 3, and 5, the best estimates we were able to make for a proved to be 0.013, 0.200, 0.200, and 0.100, respectively.

Assuming these values of a and determining new equations, we find

¹⁰ J. P. Guilford, A generalized psychophysical law, *Psychol. Rev.*, 39, 1932, 73-85.

that the exponents of S are slightly higher than those when a was assumed to be zero. Briefly, they are 1.26, 1.24, 1.29, and 1.22, respectively. The correlation coefficients are very slightly higher, suggesting improved fits to the data when a is given positive values.

The finding of positive values for the constant a calls for reflection. This means that psychological weight is zero at some *negative* value of stimulus-weight. This point is important in connection with ratio-judgments, for psychological zero is presumably the anchor about which ratio-judgments are made. One possibility is that the weight of O 's arm enters to some extent into his comparison of stimuli. The effective stimulus-weight would then be not S , but S plus some constant, assuming that the increment provided by the weight of the arm is independent of S . This is a point that calls for further investigation. Another possibility is that in making ratio-judgments, as in making other kinds of judgments, the reference point is not psychological zero but the adaptation-level. This also calls for investigation.

The type of psychophysical law for lifted-weights revealed by these ratio-judgment methods is quite different from what we would be led to expect. The positively accelerated curve is in contradiction to the negatively accelerated curve of Fechner's law. There is evidence based on results from other psychophysical scaling methods that something like Fechner's law *does* apply to lifted-weights.¹¹ We are therefore led to question whether the type of psychophysical law is a function of method. If this is the case, we may be forced to a more operational point of view than we should like, unless we can find ways of reconciling apparently discordant results.

The positive acceleration of the relation of V to S is interesting from another aspect. Positive acceleration is indicated, of course, by the fact that the exponent of S is greater than unity. If the value of m in Equation [9] is greater than unity the value of the exponent n in Equation [8] is negative. This would mean that ΔS would be a decreasing function of S , which is unheard of in psychophysical research.

The operational hypothesis may account for this, as was indicated above, but there is an additional suggestion that can be made. The exponent of S in the four equations is not very far from 1.0. This is the value of the exponent to be expected if O were to fall completely for the stimulus-error. It may be that in making ratio-judgments of lifted weights O is disposed to depend upon his knowledge of weights. There is much

¹¹ For example, Guilford and Milton Cotzin, Judgment of difficulty of simple tasks, this JOURNAL, 54, 1941, 38-52.

experience in lifting weights in association with knowledge of ounces and pounds. It is conceivable that this knowledge may function more readily when ratio-judgments are called for than when comparative interval judgments are called for. If the stimulus-error is the responsible source of the near-linear relation of R to S in these experiments, however, it is difficult to understand why the function goes beyond what would be expected; why m is greater than 1.0. One would expect some compromise by which m would be short of 1.0.

Consistency of ratio-judgments. Whatever the basis for making ratio-judgments in either the fractionation- or the constant-sum methods, our results show that there is much internal consistency. Weights that are characteristically selected to fulfill the ratio 0.5 in the fractionation-method are characteristically judged to bear the same ratio when O reports observed ratios by the constant-sum method. This was shown by Experiments 1 and 4. We can say that weights, each psychologically half as heavy as a corresponding weight in another set, are generally judged to have the same mutual ratios as their corresponding weights of double psychological quantities. This was shown by Experiments 2 and 5. Consistency is not complete, but it is surprisingly good. We can also say that half-weights scale to half as many vegs as standard weights, also shown by Experiments 2 and 5.

Evaluation of the methods. In terms of judging time and effort, the constant-sum method, either when pairs or quintets of stimuli are presented, is more economical than the fractionation-method. The derivation of scale values for particular stimuli is much more direct from constant-sum data than from fractionation-data. In fact, when all five stimuli were judged in the constant-sum method, the summed points served as immediate sources of ratio-values.

The constant-sum method offers some unique problems. One problem arose in connection with the need to extend ratios beyond sets of stimuli. In these experiments the problem was not very well met, since the establishment of common scaling between sets depended upon only one stimulus in common to the sets. The need for two sets as in these particular experiments would not often be encountered. Where it is encountered, more than one stimulus could be introduced in common, or other methods that use all of the data could be devised.

The constant-sum method promises to offer some problems of judgment that may be unique. The possible involvement with the stimulus-error has been mentioned. Examination of the data at different steps during the

derivation of ratios and *veg*-values, such as are presented in Tables III and V, shows some apparent biases that we do not have space to report here. The discrepancies between results when two versus five stimuli are presented for observation were cited in connection with the report of results for Experiment 3. There may be individual differences among the *Os* regarding the conceptions of what is meant by the instruction to divide 100 points among *k* stimuli and in the psychological operations thus generated. There seemed to be individual observer patterns of dividing points among five stimuli that carried over from the small to the large set. There seem to be other personal habits in assigning points. While the average ratios given for stimuli having the same actual psychological ratio of 2.0 in Experiment 4 were all close to 2.0, individual observer-ratios ranged all the way from 1.0 to 9.0.

Thus, the use of ratio-judgments calls for a better background of knowledge concerning their properties and a better understanding of their limitations. A general question, in this connection, is what effect the adaptation-level has on ratio-judgments.¹² We know that adaptation-level effects are common in connection with absolute and comparative judgments. When two or more stimuli are presented for ratio-judgments, are they affected in a manner that leads to a bias in the impression of their psychological ratio? These and other methodological problems concerning ratio-judgments in all ratio-methods need systematic investigation.¹³

SUMMARY AND CONCLUSIONS

The major problems of this investigation had to do with psychophysical methods utilizing ratio-judgments; specifically, the fractionation-method and the constant-sum method, and some variations and applications of the latter. Will these methods yield concordant results in terms of psychological scale values and in terms of the psychophysical law? What is the kind of psychophysical law for lifted weights to which we are led by ratio-judgment methods? The answers to these problems were sought in five experiments with lifted weights. The major conclusions are:

- (1) The fractionation- and constant-sum methods gave completely concordant ratios when stimuli were judged in pairs and the ratio judged was 2:1.
- (2) There is considerable agreement in ratios and scale-values obtained

¹² Harry Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29.

¹³ We are indebted to W. B. Michael and Andrew L. Comrey for critical reading of the manuscript, for which we accept full responsibility.

from different stimulus-levels and from different variations of the constant-sum method. There were systematic differences, however, when two versus five stimuli were presented for judgment by the constant-sum method.

(3) The same type of equation was indicated to describe the relation of psychological weight to physical weight for the three ratio-method variations employed, in four experiments.

(4) The psychophysical relationship for lifted weights seems to be a power function, where the psychological weight in veks is proportional to the stimulus-magnitude raised to some power slightly greater than 1.0. The Fechner and the Harper-Stevens laws were not supported by these ratio-judgment data. The type of relationship found, however, may be a function of the kind of psychophysical method used.

THE RELATION OF PERCEIVED SIZE TO PERCEIVED DISTANCE

By HOWARD E. GRUBER, University of Colorado

Psychologists have long been concerned with the fact that the perceived size of an object remains approximately the same even though its retinal image shrinks as distance increases. Much research devoted to the study of size-constancy may be summarized as follows: conditions of observation which may be presumed or demonstrated to impair distance-perception likewise reduce size-constancy; and, as the proximal stimulus-correlates of physical distance from *O* are eliminated, the distinction between phenomenal size and angular or perspective size vanishes.¹ These demonstrated effects upon perceived size of variations in the stimulus-correlates of physical distance have been widely taken to signify a causal relationship in which perceived size depends on perceived distance.

In studying the correlation between the perceptions of size and distance, attention must be given to the kind of distance-judgments called for. Investigators of size-constancy have frequently relied on information concerning the many variables of stimulation which are useful in the discrimination of distance. This is not surprising since most of the quantitative research on the perception of distance deals with the ability to discriminate small differences in distance. If the perceived size of an object depends on perceived distance, it must depend, however, on the total distance from *O* to the object—which is often very large in comparison with the distance between the object and its neighbors. Precise mechanisms for the discrimination of distance may lead to the correct localization of an object with respect to some neighboring object or point on a surface, but this does not tell us just what impression of distance *O* has. How far away does the object seem to him? If two objects are being compared, as in the typical size-constancy experiment, does the nearer seem one-half or one-third as far from him as the farther? The perception of the relative size of objects at different distances must depend on the effect of such *distance ratios*. Perhaps the distance-ratios or some proximal stimulus-

* Accepted for publication July 26, 1953. From a dissertation submitted to the department of psychology of Cornell University. Directed by Professor T. A. Ryan.
¹ B. E. Holaday, Die Grössenkonstanz der Sehdinge bei Variation der inneren und äusseren Wahrnehmungsbedingungen, *Arch. f. d. ges. Psychol.*, 88, 1933, 419-486; A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this JOURNAL, 54, 1941, 21-37.

correlate of it influences the perception of relative size. In any case, there is no *a priori* reason to believe that the perception of equal distance by means of zeroing adjustments involves the same mechanisms as the perception of distance-ratios.

EXPERIMENT I

The present study utilizes size-matches of objects at different distances and distance-ratio judgments of the same objects.² To test the relation of perceived size and perceived distance, two hypotheses were formulated:

Hypothesis 1. If the distance of an object is overestimated, its size will be overestimated.

Hypothesis 2. The greater the overestimation of an object's distance, the greater the overestimation of its size. This hypothesis can be tested either by measuring the correlation of judgments of size and distance in one condition of observation, or by measuring the correlation of group-trends in a series of conditions.

Observers. The *O*s were 31 volunteers from an introductory course in psychology. All had 20/20 vision or were corrected to it. None knew the purpose of the experiment.

Apparatus. The stimulus-objects were two triangles, one variable in size and displayed on a large black screen, the other variable in distance and suspended from a wire. The arrangement of the apparatus is shown in Fig. 1.

The standard distance was defined by the object which served as the size-variable, an equilateral triangle of white cardboard emerging from a horizontal slit in a black cloth screen. The screen was 107 cm. square, with an apron of black cloth continuing to the floor to conceal *E* from *O*. The exposed height of the triangle could be varied from zero to 50 cm. It could be adjusted by either *E* or *O*, as it was controlled by a pair of selsyn motors. Behind the screen a pointer and millimeter scale indicated the exposed height.

The variable distance was defined by the object which served as the size-standard, an equilateral triangle of the same white cardboard and mounted on a 20 cm. $\times$ 30 cm. cardboard rectangle covered with black cloth. It was suspended from a taut wire stretched between a point just above and behind *O*'s head and another point on the wall of the experimental room. Either *E* or *O* could control its position by means of a pulley system. A color-coded centimeter scale was painted on the upper side of the wire, invisible to *O*.

No headrest was used. *O* sat in an ordinary classroom chair with an armrest. His position was adjusted before the experiment began, and checked frequently.

Illumination was provided by two lamps, one directed at each triangle. During

² Similar reasoning can be applied to the other constancies (cf. B. K. Stavrianos, The relation of shape-perception to explicit judgments of inclination, *Arch. Psychol.*, 1945, (No. 296), 1-94). Stavrianos' experiment should be repeated with an appropriate modification.

the size-judgments, each lamp, mounted in a reflector, was 65 cm. from the object it illuminated and slightly to one side of it, from *O*'s point of view. As the lamps were fixed, their positions with respect to the triangle on the wire varied during the distance-adjustments. For high illumination, frosted 40-w. lamps were used. Although the lamps were directed at the triangles, their general illumination of the room was adequate for clear vision of the other objects and surfaces in it. For low illumination, 2-w. neon lamps were used. After the period of dark-adaptation, the

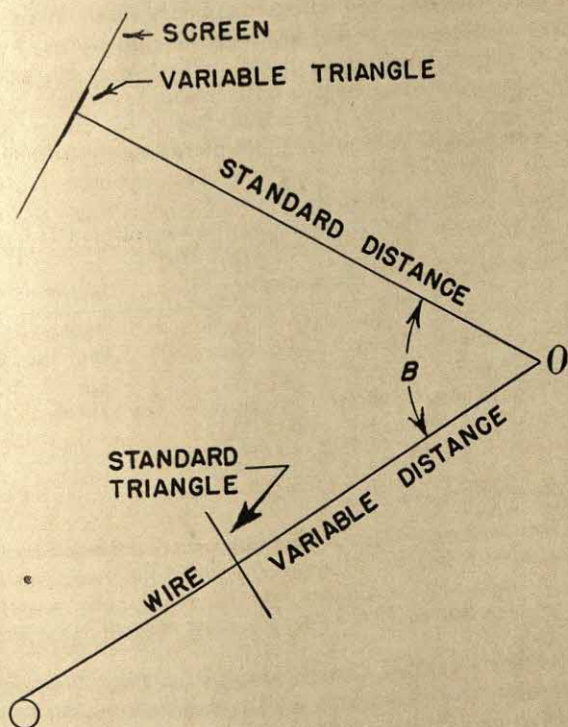


FIG. 1. POSITION OF STIMULUS-OBJECTS

In Experiment I, $B = 10^\circ$, standard triangle = 15 cm., standard distance = 330 cm.;
 In Experiment II, $B = 37^\circ$, standard triangle = 10 cm. and 15 cm. (alternated);
 standard distance = 450, 350, 250, 200, 300, 400 cm. (in the order given).

walls and floors of the room could just be perceived by virtue of faint, scattered reflections.

Procedure. Judgments were obtained first in high illumination and then, following a rest-period, in low. After *O* was seated and his position adjusted, *E* took his seat behind the screen and gave *O* brief instruction in manipulating the size-variable. This practice also served to acquaint *O* with the range of sizes available.

Judgments of equal size were obtained by the method of adjustment. With the physical distances of the standard and variable in the ratio 1:2, *O* adjusted the

variable until it appeared the same size as the standard. The instructions were designed to favor a phenomenal attitude, simply by requesting *O* to "make the triangles look the same size." If *O* raised any questions, *E* merely repeated relevant portions of the instructions.

The initial settings made by *E* before each adjustment by *O* were so predetermined that the pattern would be the same for all *O*s; the lengths of series were varied, and ascending and descending orders were alternated. Each *O* made 20 judgments in this way.

After this series, there were two further trials during which *O* was asked to give a verbal report on the process he had used in making his size-matches. If, in low illumination, *O* did not spontaneously report a difference in ease of judgment, *E* asked if there was any difference between the judgments under the two conditions.

Distance-judgments began with a verbal estimate of the distance-ratio.³ *O* was asked to "estimate the distance from you to the near triangle relative to the distance

TABLE I
SIZE-JUDGMENTS AND HALF-DISTANCE JUDGMENTS UNDER TWO
CONDITIONS OF ILLUMINATION

	Size judgment		Half-distance judgment	
	High illum.*	Low illum.†	High illum.*	Low illum.†
Correct value (cm.)	15	15	165	165
Mean (cm.)	18.35	18.37	171.8	173.7
SD	1.28	1.80	10.5	12.4
Mean percentage error	22.35	22.43	4.1	5.3
t‡	14.35	10.09	3.2	3.8

* High illumination, $N=31$.

† Low illumination, $N=30$.

‡ Difference between correct value and mean judgment; all differences significant at the 1-% level.

from you to the far triangle." If *O* answered "about one-half" he was asked to be more specific.⁴

Up to this point, the objective distance-ratio was 1:2. For the half-distance judgments, the objective ratio was variable, and the task was to determine the point at which the *subjective distance-ratio* was 1:2. *O* was asked so to adjust the distance variable that it appeared to be half as far from him as the standard. As before, the instructions called for a phenomenal attitude. After each judgment, *E* recorded *O*'s setting, and then moved it to a new starting point in preparation for the next judgment. Each *O* made 10 reports, alternating ascending and descending series with their lengths varying haphazardly.

Under low illumination the procedure was the same. After the size- and distance-

³ Between the judgments of relative perceived size and perceived distance there intervened two other types of judgment, the results of which will be reported separately. These were judgments of the limits of sizes which might possibly appear equal, and judgments of angular size.

⁴ In a later experiment performed by J. J. Gibson, O. W. Smith, and the author, a more exact psychophysical procedure was developed for obtaining verbal estimates of distance-ratio.

judgments under high illumination, there was a 10-min. rest-period, during which *O* drew pictures of persons. He was then led back to the experimental room with his eyes closed. When *E* had changed the lamps, *O* was permitted to open his eyes. The size-judgments were begun after a 5-min. period of adaptation to very dim light.

The entire experimental session, including the rest-period, lasted about 1¾ hr. Because one *O* could not complete the session, there was one *O* less in low illumination than in high.

Results: (a) Size- and distance-judgments. The results of Experiment I do not show any relation between perceived size and perceived distance. Table I presents the mean results for the size-judgments and half-distance judgments. The use of mean percentage error, rather than the Thoules or

TABLE II
RELATIONS AMONG JUDGMENTS TESTED BY CHI-SQUARE

Judgments	Illumination	N	Chi-square	P
Size	High vs. low	30	16.1	< .01
Half-distance	High vs. low	30	10.8	< .01
Half-distance vs. verbal estimate of distance ratio	High	23*	9.6	< .01
Size vs. half-distance	High	31	1.83	< .20
Size vs. half-distance	Low	30	negligible	—
Size vs. verbal estimate of distance ratio	High	23*	2.30	< .10

* Verbal estimates for 7 *O*s were eliminated as too vague.

Brunswik ratios,⁵ for the size judgment facilitates comparison with the distance judgments.

Table II presents values of Chi-square as a test of the strength of relation between different kinds of judgment. Chi-square is based on 2×2 contingency tables, dichotomized at the medians. In computing values of Chi-square for the relation between verbal estimates of distance-ratio and other judgments, only those verbal estimates were used which were definite enough to be quantified. There were only 23 such cases in high illumination and 15 in low. Since the distribution of 15 cases in a 4-cell table gives a theoretical frequency of less than 5 per cell, values of Chi-square involving verbal estimates in low illumination do not appear in the table.

It is evident from Table II that there was little or no relation between judgments of size and distance. All three values of Chi-square indicating a correlation are insignificant. One is approximately zero, and the other two are opposite in sense to the relation predicted from Hypothesis 2.

⁵ For a discussion of these ratios, see R. S. Woodworth, *Experimental Psychology*, 1938, 605.

(b) *Size-distance paradox.* The mean results, shown in Table I, held an even greater surprise. As usual in constancy experiments, the relative size of the far object was *underestimated*. That is, it was necessary to make it physically larger than the near one in order for the two to appear equal. The relative *distance* of the far object was, however, *overestimated*, i.e., when the physical distance ratio was 2:1, the far object appeared much farther than twice as far away. This constant error in

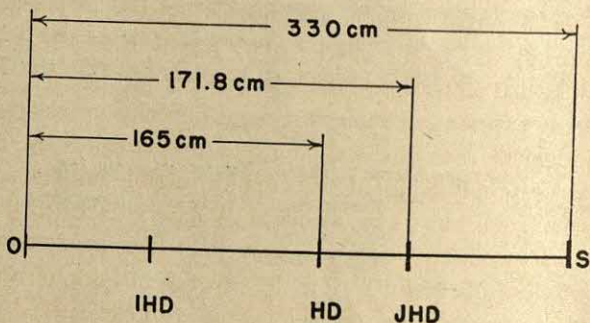


FIG. 2. THE SIZE-DISTANCE PARADOX

IHD = implicit half-distance judgment. On the hypothesis that perceived size is proportional to image size $\times$ perceived distance, if size-constancy $< 100\%$, IHD must lie between O and HD; HD = objective half-distance; JHD = judged half-distance; mean for 31 Os.

half-distance judgments can be demonstrated easily by trying to position a small object, such as a pencil, that it appears to be half way between one's eyes and some reference object (see Fig. 2).

This result directly contradicts Hypothesis 1. It may be called the 'size-distance paradox' as the error in distance perception implied in typical size-constancy results is opposite to the error actually obtained in half-distance judgments.

The results for verbal estimates of distance-ratio were in general agreement with those for half-distance judgments. This can be seen from Table II, in which the relation of the two is indicated. On the other hand, it seems that the relative overestimation of the far distance was greater in verbal estimates of distance-ratio than in half-distance adjustments. The median value of the former was 2.33, a constant error of 17%, while the median error in the latter was 8.8 cm., only 5%.

The reason for the greater constant error in the verbal estimates than in the adjustments is not entirely clear, but the fact that the verbal estimate preceded the adjustment suggests that the former was closer to O's actual

impression of relative distance during the experiment, until distance became a variable under his control.

(c) *O's reports.* The reports of the *O's* reveal that they succeeded, in general, in taking the naïve, phenomenal attitude asked for in the instructions. Few *O's* made any conscious effort to compensate for the difference in distance when making size-judgments. Occasionally, *O's* running comment indicated that he thought about the matter and avoided such an attempt, either because he realized it was contrary to the instructions or because he was uncertain as to the appropriate 'correction.' The most typical report, however, was to the effect that *O* simply looked back and forth and turned the dial until the two triangles looked equal.

O generally experienced somewhat more difficulty in making the half-distance judgment than the size-judgment. Nevertheless, he usually reported that he simply adjusted the distance variable until it appeared to be half as far from him as the standard. A few *O's* reported the use of a kind of visual geometry, with imaginary lines forming a constructed bisection, but close questioning after the experimental session showed that the actual adjustments were not compatible with the alleged operations. Probably, after the geometrical work, these *O's* made a final correction so that the distance-ratio 'looked right.'

(d) *High and low illumination.* Like judgments under high and low illumination were made almost an hour apart, during which interval *O* performed a variety of tasks and also left the experimental room for about 10 min. The degree of relation between like judgments under the two conditions, therefore, may be taken as indices of test-retest reliability. It is clear from Table II that *O's* judgments under the two conditions were quite consistent.

The differences in mean results for size- and distance-judgments in the two levels of illumination were negligible. This was unexpected, as Holaday had found an appreciable reduction in size-constancy in a similarly low illumination.⁶

Another surprise lay in the fact that the size-judgments were subjectively easier to make in the low illumination. Twenty-three *O's* reported this, while only two reported it easier in high illumination, and seven felt there was no difference. As one *O* put it, "There is less to divert you," in low illumination.

Perhaps the clearly perceptible difference in distance of the standard and variable was disconcerting while making size-judgments in high

⁶ *Op. cit.*, 424-434.

illumination, even though some stimulus-correlate of these distances must have influenced the judgment of size.

Summary and discussion. Both experimental hypotheses were contradicted. A mean overestimation of the distance of an object was coupled with a mean underestimation of its size; and there was no significant correlation of individual differences in perceived size and perceived size and perceived distance. Similar results were obtained under two radically different levels of illumination.

These results offer great difficulty for the hypothesis that perceived size is proportional to perceived distance. On the other hand, this hypothesis seems so plausible and is so widely accepted that a single experiment undermining it should be closely scrutinized.

The most important limitation of Experiment I lies in the general plan of obtaining a series of size- and half-distance judgments for only one pair of distances. It is possible that the relation of these variables is complicated by some other factor. Even if perceived size is proportional to perceived distance, the quantitative outcome of a size-match may also depend on some constant, specific to each *O* and not measured by the size- and distance-judgments obtained.

In that case, some orderly relation between size and distance would be revealed in size-matches and half-distance judgments obtained over a range of distance-pairs. For example, perhaps the trend of errors in size-judgments parallels the trend of errors in half-distance judgments for a given individual, although the two error-functions are displaced from each other in direction and amount varying with the individual. Such a result would offer an explanation of the failure to find any correlation of size and half-distance judgments at one pair of distances.⁷

An objection might also be raised to one aspect of the procedure employed in Experiment I. The screen on which the variable size-object appeared was much larger than the card which served as background for the standard. This was done to avoid the possibility of a judgment of proportionality. On the other hand, it may have introduced a contrast effect in which the variable on the screen was dwarfed by the size of its background. Individual differences in this contrast-effect may have obscured any existing correlation of size and distance.

⁷ Although the present research was completed before Gilinsky's recent paper, Experiment II, indicated by this line of reasoning, should provide a fair test of her suggestion that there is a constant which determines both size- and distance-perception, a finite limit of perceived distance, differing for each *O* in a given situation. See A. S. Gilinsky, Perceived size and distance in visual space, *Psychol. Rev.*, 58, 1951, 460-482.

EXPERIMENT II

To meet the objections raised in the preceding discussion, Experiment II was conducted. The issue now at stake was a very general prediction:

Hypothesis 3: Deviations from perfect size-constancy over a range of distances will be correlated in some way with inaccuracies in distance-ratio judgments over the same range.

In Experiment I, each *O* made a series of size- and half-distance judgments at one distance-pair. In Experiment II, each *O* made a similar series of judgments at each of six distance-pairs. In addition, two new types of judgment were obtained. These were introduced to measure any constant errors which might result from differences in the sizes of the backgrounds on which the standard and variable were displayed.

Observers. There were 20 *O*s of whom 8 were graduate students in psychology and 12 were undergraduates enrolled in an introductory course in psychology. All had 20/20 vision or were corrected to it.

Apparatus. The apparatus was the same as in Experiment I, but the room used was larger than before (see Fig. 1). Indirect and fairly homogeneous illumination was provided by four internally reflecting bulbs directed at the ceiling from points high on the walls.

Procedure. At each of six distances of the screen, four kinds of judgments were obtained in the following order: half-distance judgments ($1/2 D$); equal distance-judgments (*ED*); size-matches (*ES* $1/2$) with the standard half as far from *O* as the variable; and size-matches (*ESE*) with the standard and variable equidistant from *O*. Two standard sizes, 10 cm. and 15 cm., were alternated to minimize the possibility that *O* would become habituated to setting the variable so that its apex always appeared at the same point on the screen.

The instructions were: (1) so to set the distance-variable that it appeared to be half as far or at the same distance from *O* as the standard distance; and (2) so to set the size-variable that it appeared to be the same size as the standard. *O* was asked to avoid any 'corrections' based on knowledge, and to base his judgments on phenomenal impressions of size and distance. Before any data were taken, two practice trials for each kind of judgment were made at two distances of the screen, 400 cm. and 300 cm.

During the experiment proper, six distances of the screen were used. At each position, first $1/2 D$ and then *ED* judgments were obtained. Then *E* set the size-standard half as far from *O* as the screen and *O* made *ES* $1/2$ and then *ESE* judgments. After each judgment, *E* reset either the size- or distance-variable to a new starting position, until six judgments of a kind had been made, with alternating ascending and descending series. While the distance-judgments were being made, a triangle of 2 cm. altitude was exposed on the screen.

After each set of 24 judgments, *E* moved the screen to a new position. The experiment began with the screen at 450 cm. from *O*, this distance decreasing to 200 cm. and then increasing to 400 cm. This order permitted some estimate of practice effects while obtaining judgments at six different distances.

Results. The results of Experiment II represent another series of failures to find any relation between perceived size and perceived distance.

(a) *Group trends.* The mean judgments are given in Table III, and the mean percentage constant errors in Table IV. The size-distance paradox

TABLE III
MEAN SIZE- AND DISTANCE-JUDGMENTS AT SIX DISTANCES OF SCREEN

These values have been so adjusted that all are given in terms of a standard size of 10 cm.

($N=20$)

Distance of screen (in cm.) distance of screen	Judgment							
	$\frac{1}{2} D$		ED		ES $\frac{1}{2}$		ESE	
	M	SD	M	SD	M	SD	M	SD
450	260	16	449	16	12.3	1.2	10.6	0.8
400	238	17	400	15	11.7	1.1	10.3	0.6
350	206	15	347	14	11.5	1.8	10.5	0.7
300	179	12	296	11	11.1	1.0	10.3	0.7
250	144	10	248	7	11.0	0.9	10.4	0.7
200	115	9	199	7	10.4	1.0	10.2	.6

TABLE IV

PERCENTAGE CONSTANT ERROR OF MEAN SIZE- AND DISTANCE- JUDGMENTS

These results represent an underestimate of the relative size of the far object and an underestimate of the relative distance of the near object. In the text, the latter is treated as an overestimate of the relative distance of the far object.

($N=20$)

Distance of screen (in cm.)	Judgment			
	$\frac{1}{2} D$	ED	ES $\frac{1}{2}$	ESE
450	15.6	-0.1	23	6
400	19.0	0.0	17	3
350	17.7	-1.6	15	5
300	19.3	-1.1	11	3
250	15.2	-0.7	10	4
200	15.0	-0.5	4	2

occurred as in Experiment I, again contradicting Hypothesis 1. Moreover, Hypothesis 2 was contradicted in a new and more profound way, in that conditions of observation which produced systematic changes in size-constancy had no such effect on distance-ratio judgments. The percentage constant error for half-distance judgments ($\frac{1}{2} D$) did not vary progressively with absolute distance, fluctuating around 17%. In contrast to this, the constant error in size-constancy judgments (ES $\frac{1}{2}$) increased steadily from 4 to 24%. The frequencies with which Os increased their mean

settings of the size-variable as distance increased are significantly (5-% level) greater than would be expected by chance, as tested by Chi-square, for every interval of 100 cm.

As shown in Table III, the regular increases in variability in $\frac{1}{2} D$ and ED , as the distance of the screen increased, were not paralleled by increased variability in $ES \frac{1}{2}$ and ESE .

(b) *Correlation of individual differences.* As in Experiment I, correlation between $\frac{1}{2} D$ and $ES \frac{1}{2}$ was negligible, contradicting Hypothesis 2. Values of r for the six distances ranged from -0.10 to 0.24 . When each O 's percentage errors in $\frac{1}{2} D$ and in $ES \frac{1}{2}$ at all six distances were pooled, a plot of the pooled size-judgments against the pooled distance-ratio judgments revealed no trend. As a further test, each O 's constant errors in size- and distance-judgments, as measured by the control judgments, ESE and ED , were subtracted from $ES \frac{1}{2}$ and $\frac{1}{2} D$ respectively. A new plot, prepared after these corrections, revealed no trend.

The only hint of a correlation between judgments of size and distance was gained through a drastic process of selecting data. For each O , judgments of $\frac{1}{2} D$ were plotted against physical distance. By inspection, it was determined that these curves were quite smooth for seven O s, who can, therefore, be regarded as particularly consistent in their sets of $\frac{1}{2} D$ judgments. For these seven, the plots described in the preceding paragraph suggested some correlation of $ES \frac{1}{2}$ and $\frac{1}{2} D$ in accordance with Hypothesis 2. Even if this result can be duplicated, it would be quite limited in its application, and would hardly indicate a close causal coupling of perceived size and perceived distance.

The individual differences in size-judgments can be ascribed, in part, to effects operating independently of differences in the physical distances of the standard and variable, since correlations of ESE and $ES \frac{1}{2}$, computed for each distance of the screen, were all positive, r ranging from 0.34 to 0.71 .

The mean percentage errors in the control judgments, ESE and ED , were too low to account for the magnitude of the errors in the size-constancy and distance-ratio judgments.⁸ Of course, individual variations in small constant errors might partially obscure any correlation of these judgments, but the method of constructing a corrected plot of size-judgments against distance-judgments should have been sufficient to overcome this difficulty.

⁸ The fact that the mean error in equal distance-judgments was so nearly zero does not mean that O s were extremely accurate in this judgment. There was considerable variability among O s, but no constant error typified the group.

(c) *Analysis of individual results.* To test Hypothesis 3, a separate pair of graphs was drawn for each *O*, plotting his percentage errors in $ES \frac{1}{2}$ and in $\frac{1}{2} D$ against the distance of the screen. Inspection of the 20 pairs of curves constructed in this way yielded no confirmation of Hypothesis 3. Sometimes the two curves crossed each other once or twice, sometimes not at all. Sometimes they were convex with respect to each other, sometimes concave. In short, this method revealed no trace of a correlation of errors in size constancy and distance-ratio judgments as distance varied.⁹

Discussion. It is abundantly clear that a close causal dependence of perceived size and distance could not be demonstrated in the present experiments. Admittedly, the non-occurrence of a hypothesized correlation does not necessarily disprove the hypothesis. It might be argued, for example, that the absence of correlation in this study and the presence of it in Brunswik's indicates nothing more than two samples which differ in heterogeneity.¹⁰ To this objection three answers may be given: First, in Experiment II, as against Brunswik's, more thorough control of irrelevant variables should compensate for the greater homogeneity of the sample of responses. Moreover, the present results display considerable variability in both the size- and distance-judgments, hence the absence of correlation cannot be due to unusually homogeneous samples. Secondly, the present results offer a greater problem than the mere non-occurrence of a correlation of perceived size and perceived distance. The size-distance paradox actually *contradicts* the hypothesized relationship. Thirdly, it is possible to develop a theory which accounts for the known facts without recourse to the questionable hypothesis.

Failure to verify an important and widely accepted hypothesis imposes the task of reexamining the hypothesis itself and the facts it was intended to explain. Perhaps its original purpose was to state a functional relation between a percept and the stimuli which give rise to it. The size of the retinal image produced by an object is, of course, a dimension of proximal stimulation. The perceived distance of an object is, however, only another spatial extent, having much the same status as its perceived size. This has led to a stalemate in which students of space-perception move in a circle, sometimes explaining perceived size in terms of perceived distance, and sometimes perceived distance in terms of perceived size. The first approach, the theory of optimal stimulation, is distinguished by its emphasis on the richness of the immediate stimulus-scene out of which perceived distance

⁹ This result also contradicts Gilinsky's hypothesis (*cf.* Gilinsky, *op. cit.*, 465-466).

¹⁰ Egon Brunswik, Distal focussing of perception: size-constancy in a representative sample of situations, *Psychol. Monog.*, 56, 1944 (No. 254), 1-49, esp. 1-9.

normally arises. The second approach, the theory of impoverished stimulation, stresses the inadequacy of immediate sensory data and depends on such factors as past experience, the recognition of objects, and assumed size in accounting for perceived distance. Since the present experiments employed situations closer to optimal stimulation, the first approach will be discussed first and at greater length.

(a) *Theory of optimal stimulation.* In this view, many kinds of stimuli are available to produce a perceived world extended in distance from *O*. Naturally, *O* utilizes this strongly fortified distance-perception in a variety of ways, including the perception of size. Since the cues for distance are so diverse, it is neither necessary nor feasible to formulate hypotheses specifying the quantitative functions relating perceived size to specific *proximal stimulus-correlates* of physical distance. Rather, these correlates and other factors combine in one percept, perceived distance, on which perceived size depends. This view gains some support from the fact that several of the stimuli for distance approach a limit as physical distance increases and that, correspondingly, size-constancy sometimes decreases at similar distances.¹¹

The use of perceived distance as a catch-all for many different proximal stimuli has often led the same theorists to a *distal-stimulus-bound* orientation toward space perception. Holway and Boring, for example, conclude from an experiment in size-constancy utilizing a number of different conditions of observation, "For each of these conditions the most probable form of the function relating apparent size to distance was found to be linear."¹² The authors also conclude that "The organism can perceive the size of an object as constant, even though its distance changes, provided the perception is complex enough to provide the essential differentiae."¹³ No attempt was made, however, to isolate or analyze these 'essential differentiae.' Thus, the basic pattern of behavior in taking this approach is to refer non-committally to a group of stimuli for distance-perception, and to relate perceived size to physical distance, sometimes substituting a reference to perceived distance.¹⁴

While the foregoing is, I believe, a fair description of a widely held

¹¹ At least one recent experiment suggests that size-constancy, in some circumstances, remains approximately perfect over very long ranges of distance. J. J. Gibson (Ed.), *Motion Picture Testing and Research*, 201 ff.

¹² Holway and Boring, *op. cit.*, 36.

¹³ *Idem.*

¹⁴ Boring's intent is, perhaps, merely to postpone an admittedly necessary investigation of the proximal stimuli for distance relevant to perceived size. Brunswik (*op. cit.*, 46-47, has gone further, explicitly making the correlation of perceptual responses with distal stimuli the central subject matter of his studies in space perception.

point of view, it is also an account of the writer's own starting point in this research. The failure to demonstrate the dependence of perceived size on perceived distance led the writer to rediscover several ideas, none of them very new, which perhaps need reëmphasis and reformulation in connection with size and distance perception. (1) Each perceived property of space (such as size, distance, slant, and brightness) arises directly out of proximal stimulus-relations, and each is apprehended with a sense of immediacy. Furthermore, there is no *a priori* reason to assume that one such *percept* depends on another; indeed, the assumption leads to a logical impasse and an erroneous prediction. (2) The proximal stimulus-correlates of each perceived property of space must be experimentally isolated and an attempt made to specify the quantitative functional relation between stimulus-correlate and percept. (3) The proximal stimulus-correlates to be sought are probably themselves relations among different aspects of the total pattern of retinal arousal.¹⁵ (4) Since the investigator can only manipulate the distal stimulus-correlates, he must examine the resultant pattern of proximal stimulation closely, to determine how each of its various aspects is related to a particular property of perceived space.

In the present experiment, approximate size-constancy was maintained, and the perceived half-distance was roughly proportional to the standard distance. In that sense, obviously, perceived size and perceived distance were each influenced by some correlate of physical distance. That, however, leaves open the questions: Did *perceived* distance influence perceived size? Were both judgments influenced by the *same* stimulus correlate of physical distance? In short, were these judgments correlated with each other?

The present results revealed no correlation of perceived size and distance, suggesting that these two properties of perceived space were influenced by different proximal stimulus-correlates of physical distance. It is not suggested that this will always be the case. Size and distance are each only examples of perceived extent, and therefore both must be thought of as arising directly out of proximal stimulus-relations. Whether the relevant dimensions of stimulation will coincide or differ for perceived size and distance may well depend on the specific stimulus-situation—but a correla-

¹⁵ Wallach has demonstrated such a relation for perceived brightness, which in his results depended on the ratio of the intensities of arousal of adjacent retinal areas (Hans Wallach, Brightness constancy and the nature of achromatic colors, *J. Exper. Psychol.*, 38, 1948, 310-324). Gibson has done likewise for perceived slant, which in his results depended on the gradients of texture density in a region of retinal arousal (The perception of visual surfaces, this JOURNAL, 63, 1950, 367-384).

tion, which sometimes appears and sometimes does not, fails to indicate a close, causal correspondence.

(b) *Theory of impoverished stimulation.* Turning to situations in which the normal array of proximal stimulus-correlates is not available, it is tempting to abandon the effort to relate the perceived properties of space to variables of ongoing proximal stimulation. In the absence of any other obvious indicators of distance, it is easy to agree that perceived distance must be the effect rather than the cause of perceived size, and that perceived size is, in turn, determined by such factors as suggestion and recognition of objects.¹⁶

Hochberg and Hochberg have recently suggested that two factors, familiar size and the relative size of similar shapes, may have been confused in research on this subject.¹⁷ In their experiment, familiar size did not influence perceived distance, although the relative size of similar shapes was clearly effective. Thus, even under conditions of greatly impoverished stimulation, ongoing proximal stimulus-relations may account for perceived distance-relations.

On the other hand, this emphasis on the efficacy of ongoing stimulation should not lead us to ignore results showing that non-stimulus-factors, such as knowledge about the object have some influence on perception. We still require some theoretical scheme to cope with these intricately interwoven facts. The present research suggests one tentative step in that direction:

Perception, as a form of response to patterns of stimulation, is lawful, but not necessarily logical. At different moments, in response to the same complex matrix of sensory processes and memory traces, different percepts emerge, such as size, distance, shape, slant, and brightness. Every variety of percept may depend on various aspects of the same matrix in a different manner and degree. Thus, apparent contradictions, unacceptable in logic, are to be expected in perception: each percept emerges directly from its matrix, without any intervening intellectual process.

SUMMARY

Two experiments were conducted, with a total of 51 Os, to test the

¹⁶ This is the first and most obvious conclusion drawn from the experimental work of the Ames group, although their theoretical writings also imply that learned or assumed sizes and shapes are the most general determinants of all perceived space. See, for example, A. H. Hastorf, The influence of suggestion on the relation between stimulus size and perceived distance, *J. Psychol.*, 29, 1950, 195-217; W. H. Ittelson, Size as a cue to distance; static localization, radial motion, this JOURNAL, 64, 1951, 54-67, 188-202.

¹⁷ C. B. Hochberg and J. E. Hochberg, Familiar size and the perception of depth, *J. Psychol.*, 34, 1952, 107-114.

hypothesis that perceived size is proportional to perceived distance, when image-size is held constant. The procedure consisted of obtaining measurements of size-constancy and of perceived distance-ratio, using the same stimulus-objects for these size- and distance-judgments.

Three major facts emerged from the results: (1) An object which was consistently *underestimated* in relative size was consistently *overestimated* in relative distance. This result was labelled the 'size-distance paradox' because the constant error in distance-ratio judgments implied in size-constancy results is opposite to the error actually obtained in distance ratio judgments.

(2) Over a range of distances from 200 cm. to 450 cm., the mean constant error in relative size-judgments rose steadily from 4 to 23%, whereas the mean constant error in relative distance-judgments, rather than varying systematically with the standard distance, fluctuated around 17%.

(3) Analysis of individual differences revealed no correlation between size- and distance-judgments. Even a search for some constant of perceptual response, specific to each *O*, which might have obscured such a correlation, met with no success.

All three results are contrary to our hypothesis.

An alternative theoretical scheme was suggested in which perceived size and perceived distance do not depend on each other. Instead, both emerge from a common matrix of ongoing stimulus-processes and memory-traces; but they may be related to that matrix in different ways.

THRESHOLD FOR VISUAL FORM: FURTHER EXPERIMENTS

By JOHN KRAUSKOPF, RICHARD A. DURYEA, and M. E. BITTERMAN,
University of Texas

Some implications of a diffusion model derived from the Köhler-Wallach theory of figural after-effects¹ were examined in a recent study of threshold for visual form.² The rationale of that study and the results obtained are here briefly summarized.

Köhler and Wallach assume that a visual stimulus produces an alteration in the chemical condition of the tissue fluid surrounding the cells of the visual cortex, and that differences in potential between regions of different activity (ionic concentration) give rise to figural currents which determine perceptual experience. When one area of the retina is the site of more intense stimulation than another, a greater frequency of impulses will be conducted along the pathways originating in the more stimulated area. If in the course of their activity the nerve fibers liberate chemical substances at their termini, the rate of secretion increasing with the frequency of impulses, a pattern of ionic concentration isomorphic with the retinal image—a chemical projection of the retinal image—will be produced. It may be expected, however, that diffusion through the boundaries of the cortical figure will destroy its isomorphic relation to the retinal figure, and that the accuracy of perception will thereby be impaired. It may also be expected that the distorting effects of diffusion can be counteracted by appropriate increases in the intensity of stimulation.

Certain deductive consequences of this model were studied in an experiment on threshold for visual form. The intensity of illumination required for the perception of luminous figures briefly exposed in a dark room (form threshold) was found to vary with the ratio of perimeter to area (P/A)—a rough approximation of the relative amount of diffusion from a given figure—and with the magnitude of critical detail, although complete separation of the two determinants was not achieved. Perceptual errors (pre-threshold reports) were found to bear a striking relation to patterns of diffusion observed in a physical model. An important feature of the experiment was the employment of a method designed to get at the *perception* as contrasted with the *recognition* of form—the *O*s were not instructed concerning the forms to be presented and for each *O* only one ascending determination of the

* Accepted for publication July 20, 1953. The research here reported was conducted under terms of USAF Contract No. AF 33(616)-63. The project was administered by the Psychology Branch, Aero Medical Laboratory, Directorate of Research, Wright Air Development Center, with Dr. Kenneth T. Brown serving as Project Engineer. The authors are indebted to Dr. J. E. Hochberg for advice and criticism.

¹ Wolfgang Köhler and Hans Wallach, Figural after-effects, *Proc. Amer. Phil. Soc.*, 88, 1944, 269-357.

² M. E. Bitterman, John Krauskopf, and J. E. Hochberg, Threshold for visual form: A diffusion model, this JOURNAL, 67, 1954, 205-219.

threshold for each form was made. In the present series of experiments the dependence of form threshold on P/A and magnitude of critical detail was further investigated, and the effects of two other variables—area and exposure-time—were studied. The relation between form and brightness thresholds was examined, and the importance of the specialized method employed in the earlier experiment was evaluated.

EXPERIMENT I

The principal purpose of this experiment was to study the effect on threshold of variation in time of exposure. The diffusion model suggested that increases in time might be functionally equivalent to increases in intensity over a considerable range, and an experiment was designed to explore a portion of that range. Plans were made for the measurement of brightness thresholds as well as form thresholds to permit a systematic comparison of the two indices. This experiment, involving as it did several of the figures used in the previous study, provided an indication of the consistency of results obtained under the somewhat unusual conditions employed.

Observers. Fifty young men and women, students at the University of Texas, served in the experiment. They were untrained in psychology, and their vision, by their own reports, was either normal or corrected to normal.

Apparatus. The apparatus employed (a modified Nagel adaptometer) was the same as that of the previous experiment. Illumination was provided by three 50-w. tungsten lamps operated at 115 v. by means of a voltage-regulator. The adaptometer was equipped with three neutral density filters, each of which transmitted 5% of the incident light. To eliminate rod vision, a red filter was employed; its transmission was 5.2%, and it plotted on the spectral locus at 643 m μ . A linear diaphragm was used, to which a scale measuring percentage of maximum illumination was coupled, and a small lamp was housed in the apparatus to permit the scale to be read by E in the dark experimental room. A figure-holder, into which the stimulus-figures could be inserted individually, was fixed to the front face of the apparatus. The instrument was calibrated with a Luckeish-Taylor meter. Readings were made without the red filter and corrected for its transmission. With the diaphragm opened maximally (100 on the linear scale and 2 on the log scale), the luminance was 0.008 ft-L.

A chin-rest secured the head of O , who sat at a distance of 5 ft. from the stimulus-figures which were viewed binocularly. A tiny neon lamp, shining through a pin-hole in a small box, provided a reddish fixation-point. The box was attached to the end of a metal rod in such a way that the fixation-point was situated directly in front of the center of the figure-holder. The rod was mounted on a ball-bearing and could be rotated out of the line of sight by the action of two solenoids. The sequence of events on each trial was controlled by a bank of three interval-timers of the Hunter-Brown type.³ Following a ready-signal, E depressed a key which extinguished the fixation-

³ T. A.³ Hunter and J. S. Brown, A decade-type electronic interval timer, this JOURNAL, 62, 1949, 570-575.

light and moved it out of the field of view. After an interval of 0.5 sec., the adaptometer-lamps were turned on to present the stimulus-figure at one of five different exposure-times—0.2, 0.5, 0.8, 1.1, or 1.4 sec. After a brief interval of darkness, the fixation-light was swung back into position and turned on once more. The total time between the disappearance of the fixation-light and its reappearance was 3 sec.

Stimuli. The five forms studied in the experiment (circle, square, diamond, X and cross) are illustrated in Fig. 1. In all, nine figures were constructed—small circle, square, diamond, X, and cross (equal in area to 0.25 sq. in.) and large square, diamond, X, and cross (equal in area to 1.00 sq. in.). As may be seen from Fig. 1, square and X were 45° rotations of diamond and cross. The figure with the greatest lateral extent was the large X, which measured 1.34 in.; for the *O* seated 5 ft. away, its angular size was 1°16'. The figures were cut out of paper, and contact-negatives were made from these templates—transparent figures on opaque grounds. The



FIG. 1. THE FORMS OF EXPERIMENTS I AND II

negatives were bound to thin plates of clear plastic which could be inserted individually into the figure-holder of the adaptometer.

Procedure. Each *O* was brought into the air-conditioned dark room employed for the experiments, dark-adapted for 10 min., and appropriately instructed. Brightness and form thresholds for each figure were determined in the following manner: The adaptometer scale was set below the brightness threshold for the first exposure, and on successive trials the illumination was increased by steps of 0.05 log unit. The level at which *O* first reported light was taken as the brightness threshold, and the series terminated when *O* correctly named or described the form. Each *O* encountered each form only once during the experiment.

For each *O*, five different figures and five different exposure-times were used. For half the *O*s, the figures were large diamond, large X, small circle, small square, and small cross (Series A); for the remaining *O*s, the figures were large square, large cross, small circle, small diamond, and small X (Series B). This procedure was adopted because it was desired that each *O* should see each form only once. The dimensions of the figures in each of the two series were, of course, identical, and they were combined in the analysis of the results when it was found (as anticipated) that they yielded comparable threshold values.

Each *O* was presented with each figure at one of the five different exposure-times. Two 5 × 5 Latin squares were constructed, one for figures and one for exposure-times, and 25 different orders of presentation were derived by pairing the rows of the two squares in all possible ways. Two *O*s were assigned to each of the orders, one of the *O*s being presented with the figures of Series A and the second with the figures of Series B.

Results. The principal findings of the experiment are summarized in Fig. 2 where mean form thresholds and mean brightness (light) thresholds are plotted as a function of *P/A* (for all exposure-times combined)

and as a function of exposure-time (for all figures combined). Three separate analyses of variance were made for the independent data derived from the performance of each *O* on the first trial, and the results are presented in Table I. The first analysis (for form thresholds) showed highly significant variance due to P/A and to exposure-time, but there was no evidence of a significant interaction. The analysis for light thresholds also showed significant variance due to P/A and to time, with an insignificant interaction.

A third analysis of variance was performed to test the significance of the apparent divergence of the curves for form and light thresholds plotted as a function of P/A . The measure employed for the purpose was a

TABLE I
ANALYSIS OF VARIANCE BASED ON THE INDEPENDENT OBSERVATIONS OF EXPERIMENT I
(First trial for each *O*)

Source of variation	df	Form		Light		Form—Light	
		Mean square	F	Mean square	F	Mean square	F
P/A	4	.3466	23.84*	.1817	7.75*	.1084	4.43*
Time	4	.1762	12.12*	.1757	7.50*	.0069	0.28
$P/A \times \text{time}$	16	.0217	1.50	.0219	0.94	.0230	0.91
Residual	25	.0145		.0234		.0254	

* Significant beyond the 1% level of confidence.

difference-score obtained by subtracting the light threshold from the form threshold for each *O*. The results show a significant variance due to P/A but not to time, and no significant interaction (Table I). Apparently, form and light thresholds for all figures combined vary in much the same manner with changes in exposure-time, but form and light thresholds for all times combined are differently affected by changes in P/A . It should be noted that the mean curves for all times and for all forms which are shown in Fig. 2 are quite representative of the component functions. For example, when mean form threshold in relation to P/A is plotted separately for each exposure-time, a family of straight lines like the mean curve is obtained. When mean light threshold in relation to P/A is plotted separately for each time, a family of step-functions is obtained—in each case the first two points are found at one level and the last three points are grouped at a correspondingly higher level. Similar consistencies are found when form and light thresholds for each figure are plotted in relation to time.

Several conclusions may be drawn from these data. The linear relation between form threshold and P/A provides further support for the concept

of diffusion from which it was deduced, and the similarity of these results to those obtained in the previous experiment suggests that measurements made under the conditions employed have satisfactory reliability. The diffusion concept also makes it possible to understand why the brightness threshold does not vary with perimeter when area is constant. Loss occurs at the boundaries of the figure, and small variations in the amount of diffusion might affect the boundaries without affecting the ionic concentration at the center. The effect on the brightness threshold of area *per se* (independently of form) may perhaps be attributed to spatial summation

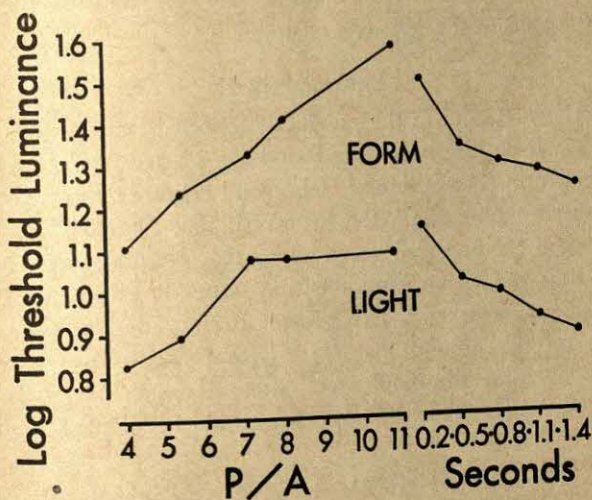


FIG. 2. FORM AND BRIGHTNESS THRESHOLDS AS A FUNCTION OF P/A AND EXPOSURE-TIME (EXPERIMENT I)

in the afferent pathway and the consequent increase in the density of concentration in the projection area.

Perhaps the most important conclusion to be drawn from the data concerns the reciprocal relationship between time and intensity. This reciprocity might be explained in terms of the conventional neurophysiological concept of temporal summation were it not for the rather extended time-intervals involved. From the standpoint of the diffusion model, which makes relatively long-term effects understandable, the reciprocity is deduced from the assumption that the rate at which ionic substances are produced by stimulation is greater than the rate at which they are dissipated. The parallel effects of time on form and brightness thresholds suggest that this interpretation is applicable at both levels.

It is rather important to note that the reciprocity between time and intensity is qualitative as well as quantitative. Although each form was seen at only one exposure-time by each *O* of Experiment I, subsequent observations on changes in form with increasing time of exposure were made by trained *Os*. For example, beginning with a time of 0.2 sec., the small cross was exposed at increasing intensities of illumination until a diamond-shaped pattern appeared. Then intensity was held constant while time was progressively increased. Under these conditions systematic changes in form were observed which corresponded to those resulting from increases in brightness—the sides of the diamond became more and more invaginated and thus the arms of the cross emerged.

EXPERIMENT II

In Experiment I, and in the previously reported experiment,⁴ consistent differences in form threshold were found for selected figures of equal area but different form. The failure of Helson and Fehrer⁵ to obtain comparable results may be explained on the hypothesis that repeated determinations with the same *Os* provided coaching in the interpretation of indistinct signs and thus yielded an experiment on recognition in which perceptual differences were masked.⁶ Experiment II was designed to test this hypothesis on the ground that if it were proved false—if significant differences in form threshold were after all obtainable with repeated determinations—more efficient procedures could be adopted in subsequent experiments.

Observers. Four advanced graduate students in experimental psychology at the University of Texas served in the experiment. All had considerable previous experience in psychophysical experimentation, and their vision was either normal or corrected to normal.

Apparatus. The apparatus of Experiment I was employed without modification in Experiment II.

Stimuli. Nine figures were used, all of which had been studied in Experiment I: large square, diamond, cross, and X (equal in area to 1.00 sq. in.) and small circle, square, diamond, cross, and X (equal in area to 0.25 sq. in.).

Procedure. With several exceptions, determinations were made in the same manner and under the same conditions as in Experiment I. In the second experiment, unlike the first, the *Os* were informed of the purpose of the research and familiar with the figures to be employed. On each trial, three thresholds were determined: the first was the light or brightness threshold; the second was the threshold for 'guessed' form—the level at which the form could be correctly inferred from minimal cues

⁴ Bitterman, Krauskopf, and Hochberg, *op. cit.*, 205-219.

⁵ Harry Helson and E. V. Fehrer, The rôle of form in perception, this JOURNAL, 44, 1932, 79-102.

⁶ Bitterman, Krauskopf, and Hochberg, *op. cit.*, 208-209.

despite the fact that it was not clearly perceived (*e.g.* a figure *perceived* as an invaginated square and *estimated* to be an X); the third was the threshold for 'certain form'—the level at which the reported form was perceived in clear, sharp outline. Each *O* served in seven sessions, the first of which was a practice session designed to familiarize him with the figures and to stabilize his criteria for the various threshold levels (particularly the second and third).

During each of the six experimental sessions, there were 10 trials. For two of the *O*s, only five figures were used in the experiment (large diamond, large X, small circle, small square, and small cross). Each figure was presented twice in each session, once at an exposure-time of 0.2 sec. and again at a time of 1.4 sec. (the shortest and longest exposure-times of Experiment I). For the remaining *O*s, all nine of the figures were used. During each session one member of each rotation-pair (*e.g.* small square and diamond) was presented at 0.2 sec. and the other at 1.4 sec.; the circle was presented twice, once at each of the exposure-times. The

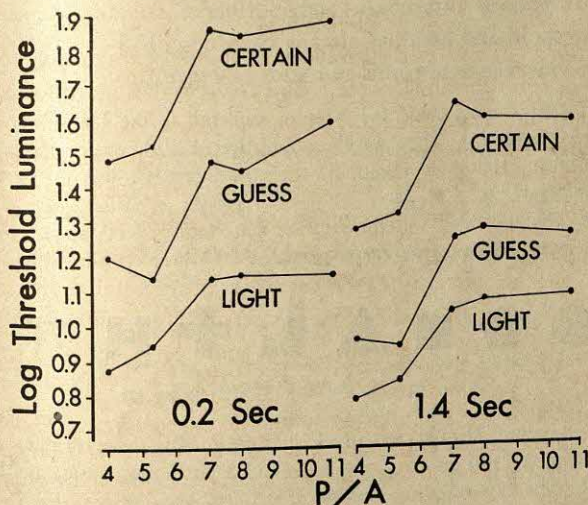


FIG. 3. RESULTS OBTAINED FROM REPEATED DETERMINATIONS BY SOPHISTICATED *O*s (EXPERIMENT II)

hypothesis was entertained that the more extensive array of figures, containing such confusion-pairs as diamond-cross and square-X, would alert the *O*s and thereby yield more precise form thresholds. In all cases the order in which the 10 determinations of each session were made was systematically randomized.

Results. The results of the experiment are graphically presented in Fig. 3. Separate curves are plotted for each of the three threshold levels at each of the two exposure-times. The results for all four *O*s were combined since the number of figures employed seemed to make no difference.

A comparison of these curves with those obtained in Experiment I clearly illustrates the effect of familiarity with the figures. In Experiment I,

a strikingly linear relationship between form threshold and P/A was found (Fig. 2), but the curves of Fig. 3 are step-functions. Under conditions of familiarity with the figures, apparently, form thresholds (like brightness thresholds) vary significantly with area but are unaffected by considerable differences in perimeter when area is constant. Presumably, as the data of Fry suggest,⁷ when the differences in form involve marked differences in the magnitude of critical detail, significant variations in form threshold may be obtained even with sophisticated *Os*. Nevertheless, it is clear that the use of repeated determinations may introduce serious artifacts.

EXPERIMENT III

The third experiment was designed to examine the effect of area on form and brightness thresholds. Some evidence for an inverse relation between threshold and area has already been described. In the present experiment, a wider range of forms and areas was investigated.

Observers. Forty young men and women, students at the University of Texas, served in the experiment. They were unsophisticated with respect to the purpose of the experiment, and their vision, by their own reports, was either normal or corrected to normal.

Apparatus. The apparatus was the same as that employed in the previous experiments, except that illumination was provided by three 100-w. lamps. By the



FIG. 4. THE FORMS OF EXPERIMENT III

calibrating procedure already described it was found that for the maximum diaphragm opening (100 on the linear scale and 2 on the log scale) the luminance was .020 ft-L. The exposure-time was set at 0.5 sec., and the viewing distance was 6 ft.

Stimuli. The eight forms studied in the experiment (circle, pentagon, square, triangle, *L*, cross, *H*, and star) are illustrated in Fig. 4. Each form was photographically reproduced (from a large paper template) in each of four areas—0.30, 0.53, 0.94, and 1.70 sq. in. The maximum lateral dimension represented in these figures was 2.5 in. (2° at a viewing distance of 6 ft.).

Procedure. Eight determinations were made for each *O*, who encountered each of the eight forms only once during the series. Two of the determinations were made at each of the four areas. The order in which the forms were presented was varied systematically according to an 8×8 Latin square, to each row of which five *Os* were assigned. Each area as well as each form appeared equally often on each trial, with combinations of area and form being randomized within the

⁷ G. A. Fry, The relation of the configuration of a brightness contrast border to its visibility, *J. Opt. Soc. Amer.*, 37, 1947, 166-175.

square. The instructions to the *O*s were the same as those given in the previous experiments. Both brightness and form thresholds were determined for each figure.

Results. The principal results of the experiment are summarized in Fig. 5, which shows form and light thresholds plotted as a function of area and as a function of mean P/A for each group of forms of the same area.

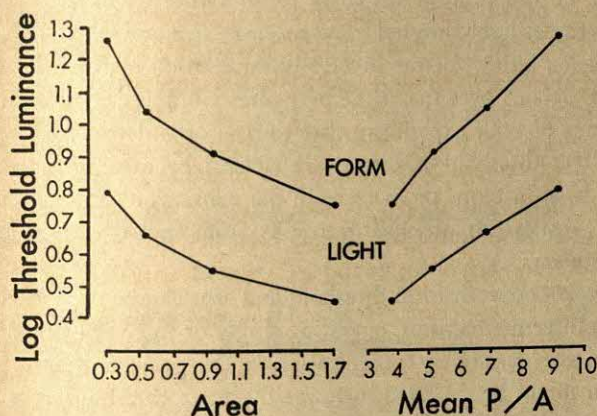


FIG. 5. FORM AND BRIGHTNESS THRESHOLDS AS A FUNCTION OF AREA AND MEAN P/A (EXPERIMENT III)

TABLE II

VARIANCE RATIO (F) FOR AREAL EFFECTS ON LIGHT AND FORM THRESHOLDS FOR EACH FORM OF EXPERIMENT III*

Form	Light threshold	Form threshold
Circle	4.70	4.49
Pentagon	5.18	34.08
Square	7.47	14.54
Triangle	4.75	7.64
Cross	3.96	15.20
L	4.48	12.32
H	11.04	18.31
Star	6.82	18.92

* For 3 and 36 df., all values shown in the table are significant beyond the 1% level except that for the light threshold of the cross, which was significant beyond the 5% level.

The significance of the differences in form threshold and the significance of the differences in light threshold attributable to differences in area were estimated separately for each form by a simple analysis of variance, the results of which are presented in Table II. Variance due to area was uniformly significant beyond the 1% level, except in the case of light threshold for cross where significance exceeded the 5% level.

The similarity of the curves for light threshold and form threshold

which are presented in Fig. 5 should not be permitted to obscure the fact that the two thresholds probably are determined by quite different mechanisms. The results of Experiment I made it clear that the decrease in the brightness threshold with increasing area cannot be accounted for in terms of decrease in the relative amount of diffusion (P/A), and the present results do not make it necessary to alter that conclusion. The linear relation between light threshold and mean P/A apparent in Fig. 5 can be regarded as an artifact of the inverse linear relation between P/A and log area which obtains over the range of figures employed in the experiment. The findings of both experiments are thus in accord with the assumption that the light threshold is directly influenced by area, and the effect is understandable in terms of increase in the density of concentration in the projection area which presumably results from spatial summation in the afferent pathways.

The relation between form threshold and area can be explained in terms of at least three mechanisms, operating singly or in combination. First, an increase in the density of concentration in the projection area by spatial summation might be expected to lower the form threshold as well as the brightness threshold; the effect in both cases would be analogous to that produced by lengthened exposure-time. Second, the linear relation between form threshold and P/A might be taken at face value—as a genuine effect of change in the relative amount of diffusion. Third, the relation might be attributed to the change in the magnitude of critical detail which is correlated with area. For the range of figures employed in the experiment there is an inverse linear relation between mean P/A and mean perimeter, which (with form constant) provides an acceptable index of the magnitude of critical detail. The separation of these three factors which is achieved in Experiment IV makes possible an evaluation of their relative importance.

EXPERIMENT IV

The fourth experiment was designed primarily to examine the rôle of critical detail. The dimensions of selected forms were systematically varied with view to sampling the effects of length, thickness, complexity, and angular relations.

Observers. Forty-eight young men and women, students at the University of Texas, served in the experiment. They were psychologically unsophisticated and their vision, by their own reports, was either normal or corrected to normal.

Apparatus. With one exception, the experimental conditions were identical with those of Experiment III: the viewing distance in Experiment IV was 5 ft.

Stimuli. The figures employed are illustrated in Fig. 6. Like those of the

previous experiments they were photographically reproduced from paper templates. The group of figures as a whole consisted of four variants of each of six basic 'forms'—cross, X, L, T, rectangle, and triangle. In Fig. 6, the stimuli (one of which appears twice) are arranged in seven series which represent several different modes of variation. The dimensions illustrated are relative; absolute values may be estimated from the length of the longest rectangle which was 1.0 in. The maximum lateral dimension in the group of figures was 1.75 in. ($1^{\circ}40'$ at a viewing distance of 5 ft.).

In the series of Xs and Ls, the length of the arms increases progressively while their thickness remains constant; as length increases, area increases markedly and P/A decreases slightly. In the series of crosses, the length of the arms increases while the area of the figures remains constant; as length increases, thickness decreases markedly and there is a considerable increase in P/A . In the series of

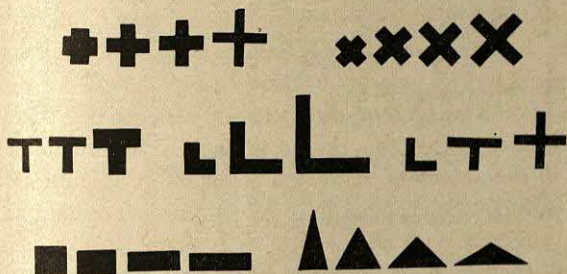


FIG. 6. THE FORMS OF EXPERIMENT IV

Ts, the length of the arms remains constant while thickness and area increase and P/A decreases. In what was designated the 'complexity' series, consisting of L, T, and cross (with the L and T being constructed by removing arms of the cross), the number of critical details increases; area increases with complexity, while P/A and magnitude of critical detail remain essentially constant. In each of these basic forms—cross, X, L, and T—the critical detail is a 90° interior angle constituted by two perpendicular arms, and the magnitude of the detail can be estimated in terms of the length of the arms.

The two remaining series consist of rectangles and triangles. In the series of rectangles, area remains constant while P/A and length increase progressively. The series of isosceles triangles of equal area was designed to study the rôle of exterior angles. The diffusion model suggests that the threshold for extremely acute and extremely obtuse angles should be high. If the threshold of the triangle as a whole is limited by that of the least readily perceived angle, the equilateral triangle should have the lowest threshold.

Procedure. Each O was shown only one variant of each of the six basic forms (making a total of six determinations). Orders of presentation were so selected that each variant of each form appeared equally often on each of the six trials. A total of 12 determinations for each variant were made for the group of 48 Os as a whole. The instructions were the same as those previously employed, with both light and form thresholds being measured for each figure.

Results. The significance of differences in threshold within each of the seven series was estimated in a separate analysis of variance, the results of which are presented in Table III. Variance due to trials was in general insignificant, both for form and brightness thresholds. Within-series variances for the brightness thresholds were, with one exception, insignificant; the

TABLE III
VARIANCE-RATIOS (*F*) FOR LIGHT AND FORM THRESHOLDS IN EACH OF THE SEVEN SERIES OF EXPERIMENT IV

Series	Light threshold		Form threshold	
	Trials	Within-series	Trials	Within-series
X	0.94	0.68	1.36	18.20*
L	0.81	2.95	1.00	17.02†
Cross	0.85	0.31	1.38	69.11*
T	2.59§	2.94	2.24	0.88
Complexity	—	1.55	—	1.53
Rectangle	2.21	3.62‡	1.17	1.50
Triangle	2.00	0.64	1.00	2.73‡

* With 3 and 39 df., significant beyond the 1-% level.

† With 2 and 28 df., significant beyond the 1-% level.

‡ With 3 and 39 df., significant beyond the 5-% level.

§ With 5 and 28 df., significant beyond the 5-% level.

variance for rectangles was significant beyond the 5-% level, but the mean values (plotted in Fig. 7) show so little order that the result may be regarded as spurious.

Inspection of the data on form threshold suggests that the principal source of variation for crosses, Xs, and Ls was magnitude of critical detail (measured in terms of the length of the arms constituting the interior angles). The validity of this conclusion is suggested by the plot of form threshold against length of arm which is shown in Fig. 7; the points constituting the three series seem to fit the same function irrespective of variations in P/A , area, and thickness. Thresholds for cross and X decrease in a similar fashion with increasing length despite the fact that, for the crosses, P/A increases markedly and area remains the same while, for the Xs, P/A decreases slightly and area increases markedly. In the series of Ts, length is constant, and there is no significant within-series variance despite simultaneous decreases in P/A and increases in area. (The mean thresholds for the Ts, which are not plotted in Fig. 7, are 1.18, 1.10, and 1.09 in order of increasing thickness.)

From these results it may be concluded that the perception of form at low levels of illumination is limited primarily by local diffusion which obscures critical detail rather than by the relative amount of diffusion from

the figure as a whole (expressed in terms of P/A). The positive relation between form threshold and P/A which appeared in the earlier experiments of this series was probably due entirely to the correlation between P/A and magnitude of critical detail for the figures employed. While the results of Experiment IV suggest that the variation of form threshold with critical detail is independent of area *per se*, it is doubtful that the results of

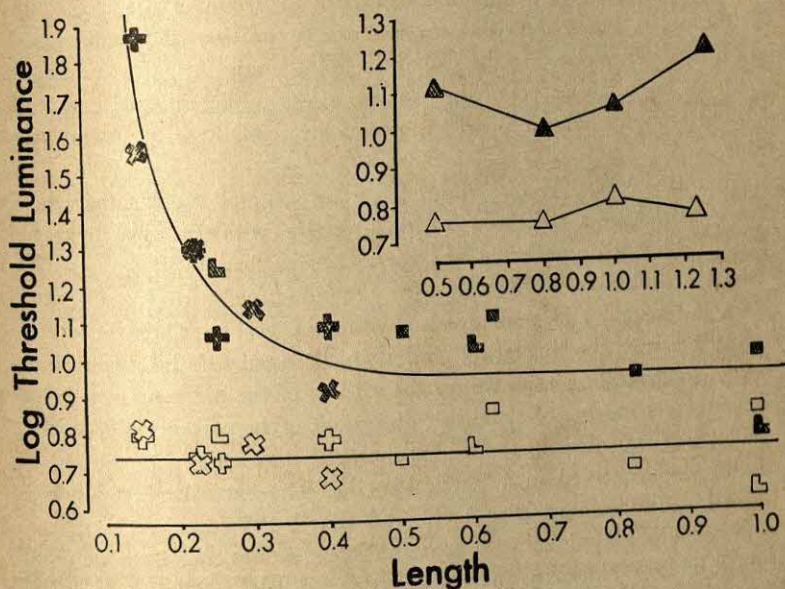


FIG. 7. FORM AND BRIGHTNESS THRESHOLDS AS A FUNCTION OF THE MAGNITUDE OF CRITICAL DETAIL (EXPERIMENT IV)

For cross, X, and L, the critical detail is an interior angle of 90° constituted by two perpendicular arms whose length provides a measure of the magnitude of the detail. The rectangles (somewhat arbitrarily) are plotted in the main portion of the figure in terms of length. In the inset, data for the series of isosceles triangles (designed to study the rôle of exterior angles) are plotted as a function of the length of base to study the rôle of exterior angles. The open points show the brightness thresholds and the filled points the form thresholds.

Experiment III can be explained completely in terms of the correlation between area and size of critical detail. In the figures of the third experiment the growth in area with linear magnitude was much greater than that in the figures of the fourth, and the plot of form threshold against linear magnitude for the third experiment shows little evidence of the sharp leveling which appears in the plot for the fourth. The data make it appear that there is some effect of area on threshold which is attributable to spatial summation and independent of the magnitude of critical detail.

For the rectangles, no significant within-series variance was found, but their lengths were such that all of the points fall on the flattened portion of the curve of Fig. 7. (The rectangles were plotted with the other figures despite some doubt that their lengths constituted an adequate measure of critical detail.) The data for the complexity series showed no significant variance in form threshold due to relatively small increases in the number of critical details. The within-series variance for triangles was statistically significant; in Fig. 7 the data are plotted as a function of length of base which is directly related to the size of the apical angle. Apparently, angular size is an important aspect of the problem of critical detail, extremely acute or obtuse angles being obscured more readily than those of intermediate value.

From the data thus far reported, certain general conclusions can be drawn: Brightness threshold varies inversely with exposure-time and with area, but it is unaffected by rather extensive changes in configuration. Form threshold varies inversely with exposure-time, with the magnitude of critical detail, and perhaps also with area *per se*. These results are compatible with the theoretical model which stimulated the investigation, but the most striking evidence for the validity of the diffusion concept lies in the correspondence of systematic distortions in the perception of shape at low intensities of illumination and at short exposure-times with the patterns of diffusion observed in a physical model. The extensive temporal and linear magnitudes more than 1 sec. and almost 1°) found to be effective in the experiments also should be noted. These values constitute the principal stumbling block to an interpretation of the results in traditional neurophysiological terms.

SUMMARY

Foveal form thresholds were measured in terms of the intensity of illumination required to identify luminous figures briefly exposed in a dark room. Form threshold varies inversely with exposure-time, with magnitude of critical detail, and to some extent with area. The reciprocity of time and intensity is qualitative as well as quantitative; with short exposure-times distortions of form appear which are comparable to those obtained at low intensities of illumination. Brightness threshold, like form threshold, varies inversely with exposure-time and with area, but it is unaffected by extensive differences in configuration. Evidence is presented to show that O's knowledge of the forms employed may mask real differences in form threshold.

VARIANCE PREFERENCES IN GAMBLING

By WARD EDWARDS, The Johns Hopkins University

For some time economists have known that the subjective value (utility) of money is different from its objective value. Recently, I have demonstrated that Ss prefer to bet at some probabilities rather than others.¹ The existence of these probability-preferences is most easily accounted for by assuming that the subjective values of probabilities are different from their objective values. Consideration of these two facts leads to a simple mathematical model for predicting gambling decisions. That model is that Ss choose among bets in such a way as to maximize *subjectively expected utility* (SEU). The SEU of a bet is calculated as follows: if U_i represents the utility of the i th possible outcome of the bet, and P^*_i represents the subjective probability of that outcome, then $SEU = \sum_i P^*_i U_i$. Experiments have shown that this model predicts choices among simple bets reasonably well.²

Allais and others have hypothesized that people may not only prefer to bet at some probabilities rather than others, but may also prefer to bet at some variances rather than others.³ In other words, some people might prefer bets with high possi-

* Accepted for prior publication August 1, 1954. This experiment was done under Contract N5-ori-166, Task Order I, between the Office of Naval Research and The Johns Hopkins University. This is Report No. 166-I-186, Project Designation No. NR 145089 under that contract. Preliminary Experiment I was carried out in the Psychological Laboratories of Harvard University and was partially supported by funds made available from the Laboratory of Social Relations, Harvard University, through its Project on the Nature of Risk and the Theory of Games. The report of Preliminary Experiment I is adapted from part of a doctoral dissertation submitted to Harvard University. I am grateful to Mrs. Harriott Quin and Mr. Klaus Rennert for collecting and processing the data of Preliminary Experiment II and the Main Experiment.

¹ Ward Edwards, Probability-preferences in gambling, this JOURNAL, 66, 1953, 349-364. Throughout the present paper, this experiment will be referred to as the first probability-preference experiment and will not be separately cited. See also Edwards, Probability-preferences among bets with differing expected values, this JOURNAL, 67, 1954, 56-67; and Edwards, The reliability of probability-preferences, this JOURNAL, 67, 1954, 68-95.

² For a full discussion of the concept of utility, the concept of subjective probability, and the SEU model, see Edwards, The theory of decision making, *Psychol. Bull.*, 51, 1954, 380-417. For a report of an experiment testing the SEU model, see Edwards, An attempt to predict gambling decisions, *Proceedings of the Dunlap Symposium on Mathematical Models in the Social Sciences* (in press).

³ Maurice Allais, Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine. *Econometrica*, 21, 1953, 503-546. Also Allais, La psychologie de l'homme rationnel devant le risque: La théorie et l'expérience. *J. Soc. Statist.*, Paris, 94, 1953, 47-73.

bilities of winning or losing, while others might prefer bets with smaller possibilities of winning or losing. It is possible, of course, to derive such preferences from appropriately chosen utility or subjective probability-scales, and such a possibility is therefore not disturbing, but such preferences might occur even when consideration of utility and subjective probability alone would not predict them. This, if true, would be disastrous for the *SEU* model discussed above.

It is necessary to be explicit about the differences among variance-preferences, utilities, and subjective probabilities. If we ignore subjective probabilities for a moment, we can conceive of an expected utility for a given bet. This expected utility will have a variance associated with it. The size of this variance will depend on the probabilities involved and the amounts of money involved. If we introduce subjective probabilities into this scheme, they will change the size of the expected utility (now a *SEU*) and of the variance, but the variance is still a number, dependent on but different from the utilities and subjective probabilities involved.

PROBLEM

The hypothesis under consideration is that *Ss* will prefer to bet at some variances rather than others. Testing this hypothesis experimentally is a difficult problem. It is impossible to manipulate variance as an independent variable without simultaneously manipulating either utility or subjective probability or both. Consequently, it is impossible to make rigorous experimental distinctions among these three factors. Instead, it is necessary to find experimental situations in which positive results can be more parsimoniously explained by means of variance preferences than by means of subjective probability or utility functions. This paper will report two preliminary experiments and one main experiment aimed toward that end.

METHOD

Apparatus. The gambling in these experiments was done on a rigged pinball machine which has been fully described in the report of the first probability-preference experiment. There are eight cells at the bottom of the machine, into any one of which the ball may roll. The *Ss* are told that the ball is just as likely to roll into any one cell as into any other. Actually, the machine is controlled by concealed electromagnets and a programmer. The program was completely random in the experiments reported here unless otherwise stated.

Procedure. The *Ss* were first shown the pinball machine, told about the bets, and told that they might either win or lose, but that the situation was biased in their favor by \$1 per session. This bias was achieved by giving each *S* \$21 in poker chips at the beginning of each session and requiring him to return only \$20 in chips or in cash at the end of the session. The *Ss* knew that if they had less than \$20 in chips, they were required to pay the difference in cash from their own

funds, while they were paid in cash any excess over \$20. After each S had finished all experimental sessions, he participated in one extra session. The sole purpose of this session was to cause him to win or lose enough money so that he won an average of \$1.00 per hour for the time he served in the experiment.

PRELIMINARY EXPERIMENT I

Subjects. The Ss (11 men) were randomly selected from among the undergraduates of Harvard College. They had previously been through the first probability-preference experiment, which used the same pinball machine and general procedure, and similar bets.

Bets. Five bets were used. The smallest one read: "If you roll a 2, a 4, a 6, or an 8, you win \$1.00. If you roll anything else, you lose \$1.00." The other four were identical except that they involved possible wins or losses of \$2, \$3, \$4, and \$5. These five bets were exhaustively paired with one another according to Ross's optimal order for pairs in the method of paired comparisons.⁴

Procedure. The cards containing these pairs of bets were mixed in with other cards used for other experiments.⁵ There were four sessions, in each of which every S chose one member of each pair once.

Results. The main result of this experiment was that choices differed from one individual to another. No consistent features of the pattern of choices from person to person could be found. Every S, however, was consistent with himself; that is, he chose the same things on successive days to the same degree found in the previous experiments. These individual patterns of choices could be accounted for by means of variance preferences—but they could also have been predicted from suitably chosen and reasonable utility curves, without calling on the notion of variance preferences at all.

PRELIMINARY EXPERIMENT II

The results of Preliminary Experiment I indicated that, if variances were important in determining choices among bets, it would be necessary to make Ss prefer high variances at one time and low variances at another time. In other words, it would be necessary to demonstrate that Ss would, given an appropriate situation, adopt a strategy which centers on maximizing or minimizing variance. This experiment was undertaken to try out a device for inducing that strategy.

Subjects. The Ss (12 men) were randomly selected from among the undergraduates of Johns Hopkins University. They had previously served in a replication of the original probability-preference experiment and in a military-game situation.⁶

Bets. The bets, 24 in number, were made with probabilities of winning running from 1/8 through 8/8 at three different expected value levels: + \$0.525, - \$0.525, and \$0.00. For complete table of them, see the first probability-preference experiment.

Procedure. S was run as usual except for a change in the rules for the payoff. He was now given \$6.00 if he won \$3.00 or more in chips and was required to pay

⁴R. T. Ross, Optimum orders for the presentation of pairs in the method of paired comparisons, *J. Educ. Psychol.*, 25, 1934, 375-382.

⁵Edwards, The reliability of probability-preferences, *loc. cit.*

⁶*Ibid.*

\$6.00 if he lost \$3.00 or more in chips; no money changed hands if he came out between plus and minus \$3.00 in chips. There were four sessions.

Results. The purpose of the special payoff arrangement was to provide Ss with inducements to use strategies based on variance differences. If S adopted the obvious strategy, he would choose bets with low probabilities of winning but high potential wins (*i.e.* he would try to maximize his variance) if he was behind in chips by \$3.00 or more. Similarly, he would try to minimize his variance if he was ahead by \$3.00 or more. Actually, that strategy was not necessarily the best available to him, especially if he knew that he had only a few bets to go in a day's session. It is, then, somewhat debatable what optimal strategy would have been toward the end of each session. The question, however, is irrelevant, because the results showed that no strategy involving variance was used. All the choices could be attributed to probability-preferences.

MAIN EXPERIMENT

The difficulty with the first preliminary experiment was that it involved no particular inducements for Ss to pay attention to variances, and the results were too easily attributable to utility functions. The difficulty with the second preliminary experiment was that it was not sufficiently sensitive to the effects of variance preferences, pre-existent or induced, if such were present. The main experiment attempted to deal with both of these problems.

The basic notion underlying the main experiment was that if variances are important in determining how Ss choose among bets, then it should be possible to demonstrate this fact by arranging situations in which it was desirable to prefer bets of either high or low variance. The method of payoff described in Preliminary Experiment II was used for this purpose. In the main experiment, however, the program for the pinball machine was so arranged that in one part of the experiment Ss were heavy winners, hence it was to their advantage to choose bets of low variance; in another part of the experiment they were heavy losers, hence it was to their advantage to choose bets of high variance.

Subjects. The Ss (12 men) were randomly selected from among the undergraduates of Johns Hopkins University. They were naïve about such experiments.

Bets. The experiment was divided into five divisions of two sessions each. In Sessions 1 and 2, the bets used in the first probability-preference experiment were again used. This was a replication of the first probability-preference experiment, intended to establish the comparability of this sample of Ss to other similar samples. In Sessions 3-10, a new set of cards was prepared. Two lists of bets of zero expected value were used: a list of high variance and one of low. Table I shows these lists. Every bet in each list was paired with all the bets in the other list. This should have resulted in 64 pairs, but actually the 8-8 pair was eliminated because the two bets are identical, hence there were only 63. Every pair was typed

on 3×5 in. cards, the cards were shuffled, and once through the deck completed a session.

Procedure. Except as described below, the procedure was the same as used before. In Sessions 1-4, S received what he won or paid what he lost in cash after every session. In Sessions 5-10, he was paid \$6 if he won \$3 or more in chips and he paid \$6 if he lost \$3 or more in chips. Sessions 3-4 were conducted to familiarize S with the new bets, and the results from these sessions are not reported. In Sessions 5-10, the pinball machine was rigged. In Sessions 5-6, it was so set that S would remain between plus and minus \$3 in chips throughout the sessions; in Sessions 7-8, that S won heavily at the beginning of the sessions (considerably more than

TABLE I
HIGH VARIANCE BETS*

- (1) If you roll a 4, you win \$4.20. If you roll anything else, you lose \$0.60. ($\sigma^2=2.52$)
- (2) If you roll a 1 or a 7, you win \$2.10. If you roll anything else, you lose \$0.70. ($\sigma^2=1.47$).
- (3) If you roll a 2, a 4, or a 6, you win \$1.40. If you roll anything else, you lose \$0.84 ($\sigma^2=1.18$)
- (4) If you roll a 2, a 4, a 7, or an 8, you win \$1.05. If you roll anything else, you lose \$1.05. ($\sigma^2=1.10$)
- (5) If you roll a 2, a 3, a 5, a 7, or an 8, you win \$0.84. If you roll anything else, you lose \$1.40. ($\sigma^2=1.18$)
- (6) If you roll anything but a 3 or a 6, you win \$0.70. If you roll a 3 or a 6, you lose \$2.10. ($\sigma^2=1.47$)
- (7) If you roll anything but a 5, you win \$0.60. If you roll a 5, you lose \$4.20. ($\sigma^2=2.52$)
- (8) Regardless of what you roll, you neither win nor lose. ($\sigma^2=0$)

LOW VARIANCE BETS*

The eight low variance bets are identical with the eight high variance bets listed above except that each sum of money is only half as large. The variance of each low variance bet is one-fourth the variance of the corresponding high variance bet.

* The variances given above are not, of course, the variances of the SEUs of the bets. The SEU variances could not be determined without knowledge of the individual subjective probability and utility curves, and would be different from S to S. Each variance given above is the variance of the objective expected value of the bet it follows. If A and B are the two possible outcomes, then the objective variance of any two-alternative bet is $\sigma^2=pq(A-B)^2$.

\$3) and stayed ahead thereafter; and in Sessions 9-10, that S lost heavily (considerably more than \$3) and stayed behind thereafter. Thus, in Sessions 5-6, S should have chosen normally; in Sessions 7-8, he should have tried to minimize his variance; and in Sessions 9-10 he should have tried to maximize his variance.

Results. The design of this experiment violates the independence and normality assumptions necessary for analysis of variance. No *F*-tests can be made. It is possible, however, to arrange the results in an analysis of variance matrix and to compute sums of squares. In such an analysis, the sources of variance are: Ss; sessions (5-6, 7-8, and 9-10); bets of high versus low variance; and probabilities. The entry in a single cell in the matrix is the number of times a particular bet was chosen over all others with which it was compared by one S in one pair of sessions. The sums of squares and the percentage which each is of the total sum of squares are

given in Table II. Since the design of the experiment makes it impossible to observe any sum of squares attributable to *S* and conditions of payoff, it is in their interactions with other variables that the effects of these variables should be sought. The most important sources of variance are clearly the probabilities and their interaction with *S*s. Next comes the interaction of *S*s with lists of bets. Nothing else is of any size. So far, then, probabilities and *S*s seem to be very important in determining choices, and variances (lists) interact with *S*s. It is interesting that there is essentially no inter-

TABLE II
SUMS OF SQUARES ATTRIBUTABLE TO EXPERIMENTAL VARIABLES

Source	Sum of squares	Percentage of total sum of squares
<i>S</i> s	0*	0
Payoff conditions (sessions)	0*	0
Variances (lists of bets)	169	1.5
Probabilities	2577	22.7
Subjects $\times$ sessions	0*	0
Subjects $\times$ lists	2243	19.7
Subjects $\times$ probabilities	3897	34.3
Other simple interactions	543	4.8
All higher order interactions	1933	17.0†
Total	11362	100

* The design of the experiment requires that these sums of squares be zero.

† All the higher-order interactions were calculated. None is individually larger than 7% of the total.

action of lists with conditions of payoff. This is in conflict with the original hypothesis that *S*s should choose bets of low variance in Sessions 7-8 and of high variance in Sessions 9-10.

Table III shows the mean number of high and low variance choices made in Sessions 5-6, 7-8, and 9-10.⁷ It is immediately apparent that all effects are small, but in the expected directions. In Sessions 7-8, *S*s chose fewer bets of high variance; in Sessions 9-10, they chose more. The difference between Sessions 5-6 and 7-8 is insignificant by the *t*-test; difference between Sessions 5-6 and 9-10 is significant beyond the 0.02-level. (All *t*-tests reported here are two-tailed; however, the use of a one-tailed test would not make the insignificant tests significant at the 0.05-

⁷ Bet 8 of the high variance list is not a high variance bet; it is lower in variance than any of the bets from the low variance list with which it is compared. In principle, then, it should not be included in scoring the results of the high variance bets. These results have been scored both ways, and the pattern of results and significance tests come out exactly the same way. Consequently, it is included in the high variance list in those results reported here.

level.) It might be argued that the effects of the different arrangements of payoff on choice of bets with high or low variance should not show itself until after *S*'s status of winning or losing is well established and this might account for the insignificant results in the case of the winning sessions. To check this argument, the individual winning-and-losing records were examined. It was found that if only Responses 33-63 were considered, then all but two of the *S*s would be ahead more than \$3.00 in chips in Sessions 7-8 and behind the same amount in Sessions 9-10. (Complete control over the amount of winning and losing during each session was not possible, since *S*'s choices could not be anticipated. All we could do was to prepare programs such that the choices we expected would win or lose heavily. A few *S*s failed to follow our expectations.) The numbers of

TABLE III
MEAN NUMBERS OF CHOICES OF BETS OF HIGH AND LOW VARIANCE

Variance	Sessions		
	5-6	7-8	9-10
High	28.5	26.5	32.9
Low	34.5	36.5	30.1

choices high and low variance bets among Responses 33-63 of each session were, therefore, tabulated. The pattern was the same as in Table III and the results of the *t*-tests were unchanged, except that the level of significance of the comparison between Sessions 5-6 and 9-10 rose to 0.001. The difference between the percentages of choices of high variance in Sessions 5-6 when all responses are considered and when only Responses 33-63 are considered is insignificant. The same is true for Sessions 7-8. The corresponding difference for Sessions 9-10 is significant beyond the 0.02-level in the expected direction.

So far the results show that only very small (though in some cases significant) effects of the experimental conditions on variance preferences can be found. A picture of variance preferences uncontaminated by probability preferences might, however, show more effects. Fortunately, this picture can be obtained from the data. In every session there were seven choices between two bets with identical probabilities but different amounts of money (and, therefore, different variances). Table IV shows the total numbers of such choices for every *S* for the different sessions. The total number of bets of high variance chosen is somewhat greater in the winning sessions and considerably greater in the losing than in the control sessions. The increase in bets of high variance from Sessions 5-6 to 7-8 is

inconsistent with prediction and with the data when all bets are considered. It is not, however, statistically significant. The increase from Sessions 7-8 to 9-10 is significant beyond the 0.01-level.

The totals at the bottom of Table IV are somewhat misleading. The main fact to be learned from this table about Sessions 5-6 and 7-8 is that most of the Ss preferred bets of either high or low variance rather consistently when probabilities are held constant. (This finding is exactly analogous to the results of Preliminary Experiment I.) A total Chi-square testing whether these individual choices are significantly different from

TABLE IV
NUMBERS OF CHOICES OF BETS OF HIGH AND LOW VARIANCE FROM PAIRS OF BETS
WITH EQUAL PROBABILITIES WITHIN EACH PAIR

S	Sessions and Variances					
	5-6		7-8		9-10	
	high	low	high	low	high	low
1	0	14	0	14	3	11
2	0	14	2	12	5	9
3	1	13	3	11	7	7
4	14	0	14	0	14	0
5	12	2	14	0	11	3
6	7	7	9	5	10	4
7	3	11	6	8	8	6
8	0	14	2	12	7	7
9	1	13	0	14	5	9
10	8	6	3	11	4	10
11	0	14	0	14	6	8
12	2	12	2	12	8	6
Total	48	120	55	113	88	80

50-50 comes out at 110 for Sessions 5-6 and 102 for Sessions 7-8, both of which with 12 d.f. are significant way beyond the 0.001-level. In Sessions 9-10, the Ss were much less consistent in their preference of high or low variance bets. The total Chi-square is only 32, which is significant at the 0.01-level. The evidence of this analysis, then, is that Ss did have preferences for one variance over another when the factor of probability preferences was eliminated from their choices, that these preferences differed from S to S, and that they were affected by the traumatic experience of frequent loss in Sessions 9-10. The nature of this effect is an increase in preference of high variance bets, as predicted.

Probability-preferences. All the evidence shows that the most important variable in determining choices in this experiment was that of probabilities. In the first two sessions, when all Ss were run on the old probability-preference bets, they showed patterns indistinguishable from those obtained

in the original probability-preference experiment and its replication. It remains to inquire whether that pattern can be recovered from the data of Sessions 5—0. Fig. 1 shows the number of times each bet of high variance was chosen over all others with which it was compared, divided by the total number of choices of bets of high variance; Fig. 2 shows the same thing for bets of low variance. It is evident in both figures that the pattern of probability-preference made familiar in previous experiments emerges from this one, even though the kinds of choices available in this experiment are

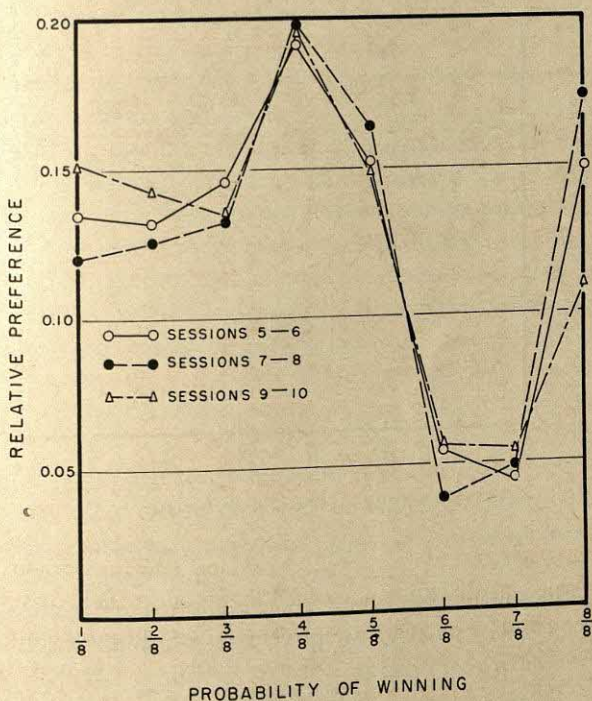


FIG. 1. PROBABILITY-PREFERENCES FOR BETS OF HIGH VARIANCE

so different. The same peak at 4/8 probability of winning, and valley at 6/8, appear. The only anomaly is the valley at 7/8, and even that valley appeared in the results of the replication of the original probability-preference experiment, when Johns Hopkins students were used as Ss.⁸

Examination of individual records makes it even clearer that probability-preferences were the dominant factors in determining choices. For instance,

⁸ Edwards, The reliability of probability-preferences, *loc. cit.*

the choices between bets with equal probabilities account for about one-third of the difference between the bets of high and low variance even though they are only one-ninth of the total number of choices. In short,

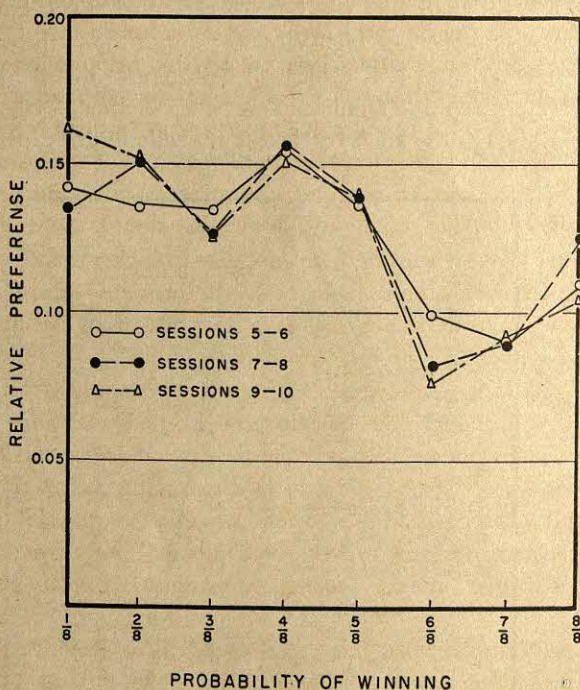


FIG. 2. PROBABILITY-PREFERENCES FOR BETS OF LOW VARIANCE

when the factor of probability-preferences is allowed to operate, it strongly overrides the effect of any variance preferences which may exist.

DISCUSSION

The main finding of this experiment is that even in circumstances where the variances of bets are supposed to be especially important, probability-preferences are the main determiners of choices among bets. Small effects of conditions of payoff on the numbers of choices of high and low variance bets can be demonstrated, and some of these effects are statistically significant, but they are very small relative to the size of the effects of probability-preferences.

When the effects of probability-preferences are held constant by considering only choices between pairs of bets having the same probabilities,

the existence of variance preferences can, however, be more definitely demonstrated. Some Ss prefer bets of high variance; others of low variance. These preferences are rather sharp and are not correlated with the probabilities or amounts of money involved. An attempt to construct utility curves which would predict them for several Ss indicated that it cannot be done; the necessary utility curves have entirely too many inflection-points to be reasonable. These variance preferences are not much affected by the winning program applied to the pinball machine in Sessions 7-8; but they are substantially disrupted by the losing program applied in Sessions 9-10. The nature of the effect is, as strategic considerations would predict, an increase in choices of bets of high variance in Sessions 9-10. Examination of total choices shows an insignificant decrease in high variance choices in Sessions 7-8, but choices only among bets with the same probabilities show an insignificant increase. In any case, it is clear that losing has much more drastic effects on variance preferences than winning.

How do these variance preferences affect the subjectively expected utility model discussed in the introduction? It seems likely that they are relatively unimportant. The evidence of this experiment strongly suggests that variance preferences are a secondary factor in determining choices among bets, important only when probability preferences (subjective probabilities) have been held constant. It is an experimental question, not answered in this experiment, whether variance preferences are also secondary relative to utilities.

A purist can complain that this experiment fails to demonstrate conclusively that the effects which are being attributed to variance preferences should not be attributed to utilities instead. This is true. An attempt to construct utility curves which would predict choices among bets with the same probabilities resulted, however, in very implausible utility curves. It is very difficult, furthermore, to see why the difference between the various payoffs should result in differences in the utility curves which would lead to the observed differences in choices, while it is very easy to see why the differences in payoff should lead to changes in variance preferences. In short, although variance preferences are not the only possible explanation of the results of this experiment, they seem to be the most parsimonious.

SUMMARY

Previous experiments suggest a mathematical model for predicting decisions among bets. This model asserts that people so choose among bets as to maximize subjectively expected utility. The possibility exists, however, that people also prefer some bets to others on the basis of their variances,

in addition to the utilities and subjective probabilities involved. This, if so, would be damaging to the model.

Two preliminary experiments failed to demonstrate the existence of variance preferences. In the main experiment, Ss were required to choose between bets with high and low variances. The most significant variable determining choices was the probabilities involved, and the probability preferences found were essentially the same as those found in other experiments. When probabilities were held constant, however, the choices did reveal the existence of variance preferences. The introduction of conditions making high variance desirable produced more bets of high variance, but the introduction of conditions making low variance desirable did not produce more bets of low variance. This is due to the fact that Ss tend to change their habits when they lose money, but not when they win.

It is concluded that variance preferences exist, but they are secondary in importance as compared with probability preferences.

TRANSFER AND GENERALIZATION OF THE INHIBITORY POTENTIAL DEVELOPED IN ROTE SERIAL LEARNING

By T. G. ANDREWS, SOLOMON SHAPIRO, and CHARLES N. COFER,
University of Maryland

There is a large body of knowledge about the positive effects in transfer of training, but there is relatively less information available regarding negative or interference-factors in transfer. Quite apart from the experimental problems of transfer of training, there has appeared a series of investigations following the formulations of Hull and his associates regarding the development of inhibitory potentials during the learning process.¹ Such postulated negative potentials may serve to describe certain situations in the transfer of training in which negative effects are obtained. Also the concept of negative potential may be useful in experiments designed to analyze the possibilities of true negative transfer, as studied by Bills and McTeer.²

The purpose of the present investigation was to determine whether the negative potential, assumed to develop in rote serial learning, would transfer to identical and to similar stimulus-items in subsequent paired-associates learning. The inhibitory potential has been postulated in part to account for the phenomenon found in serial position, as graphically described by the bow-shaped curve of serial position, and it has been useful in offering an explanation for the differential ease of learning the items in a serial list. Hull and his associates have referred to the inhibitory effects as resulting from the presence of an inhibitory potential, which they have related to an assumed principle of 'inhibition of delay.'³ It is of theoretical interest to determine whether the inhibitory effects would carry over to situations other than those in which they were developed.

Other explanations have been offered for the intraserial differential difficulty.⁴ Hull's formulation regarding inhibitory potential is used in

* Accepted for publication July 30, 1953.

¹ C. L. Hull, C. I. Hovland, R. T. Ross, M. Hall, D. T. Perkins, and F. R. Fitch, *Mathematico-Deductive Theory of Rote Learning*, 1940, 1-329.

² A. G. Bills and W. McTeer, Transfer of fatigue and identical elements, *J. Exper. Psychol.*, 15, 1932, 23-36.

³ Hull, et al., *op. cit.*

⁴ M. Foucault, Les inhibitions internes de fixation, *Année Psychol.*, 29, 1928, 92-112; J. A. McGeoch, *The Psychology of Human Learning: An Introduction*, 1942, 112.

the present investigation because of its implications for transfer of training and its position in modern learning theory.⁵

The general situation was arranged to have *S* learn a series of nonsense syllables and then learn a set of paired-associates in which the stimulus-items in the pairs were taken from the original serial list and were to be associated with new response-items. Within this general experimental situation, the major hypotheses tested were:

(1) Those syllables learned near the middle of the serial list will be more difficult to learn in immediately subsequent presentation by paired-associates. This hypothesis predicts the transfer of inhibitory potential.

(2) The more central syllables in the serial list will be relatively more difficult to learn than those near the ends when learning the paired-associates follows immediately after the serial learning as contrasted with the effects of a longer intervening delay between the two learning experiences. This prediction is based on the supposed rapid dissipation considered as a characteristic of the inhibitory potential.

(3) Syllables which are either phonetically or physically similar in appearance will also demonstrate comparable difficulties in the learning of a second set of verbal materials. This hypothesis is important in relation to transfer studies in general, in that it involves investigation of the extent of generalization of the transferred inhibitory potential to similar items.

METHOD AND PROCEDURE

Subjects. The *Ss* were 48 elementary school children selected as homogeneous and superior with regard to measures of *IQ*, reading ability, and teacher reports of 'coöperativeness' and 'interest.' The age-range was 9 to 13 yr. The *Ss* were randomly assigned to the experimental conditions described below, but the assignments were controlled to allocate an equal number from each sex to each experimental group.

Apparatus. The same apparatus was used for both the serial and the paired-associates learning. It consisted of a vertical frame with a $4\frac{1}{2} \times 13\frac{1}{2}$ -in. aperture, the right half of which was covered by a hinged plate which could be raised manually by a lever to expose this side. Slots in the rear of the frame were arranged to hold a series of cards, $7 \times 13\frac{3}{4}$ -in. in size. Only the left side of the aperture was used for presenting the serial learning lists. The right side was exposed to present the response syllable in learning by paired-associates.

The syllables were printed in black letters $\frac{1}{2}$ in. in height in the centers of the exposed portions of the white cards. Leroy Lettering Set templates aided to insure uniformity of the printed materials. An electronic metronome was used by *E* in timing the presentation of the stimulus-cards.

Experimental design and structure of the stimulus-cards. The *Ss* all worked

⁵ Hull, The conflicting psychologies of learning—a way out, *Psychol. Rev.*, 42, 1935, 491-516; Hull, et al., *op. cit.*

according to the following outline: practice learning period, serial learning, paired-associates learning. Five pairs of letter-number combinations were learned in the practice stage, and these were the same for all Ss. The material used in the serial learning and the final paired-associates learning is shown in Table I. Twenty-six of the 32 nonsense syllables in Table I have appeared in previously published lists of nonsense syllables having low association value.⁶ The remaining six syllables needed were approved by six judges as probably not having high associative value. In accordance with Hovland's procedure, the syllables in this experiment were pronounced and not spelled aloud.⁷

TABLE I
NONSENSE SYLLABLES USED IN SERIAL AND PAIRED-ASSOCIATES LEARNING

Stimulus-members of paired-associates			Changes in original syllables to produce related syllables		Paired-associate response member
Original*	Related				
condition A & D	condition B & E homo- phones	condition C & F phys. sim.	condition B & E homo- phones	condition C & F phys. sim.	
FEX	FAX†	PEM†	vowel	both consonants	BUV†
TOJ†	TUJ†	MOG	vowel	both consonants	YAD†
PEJ†	DEJ†	JEP†	1st. cons.	cons. reversed	WOB†
BIK	BIQ†	KIB†	2nd. cons.	cons. reversed	PAF†
MIB†	MIP†	MIF†	2nd. cons.	2nd. cons.	KEV†
VOF†	VOV†	VUP†	2nd. cons.	vowel & 2nd. cons.	TAH†
ZET†	ZED†	ZAL†	2nd. cons.	vowel & 2nd. cons.	NIZ†
COZ	KOZ	CAX†	1st. cons.	vowel & 2nd. cons.	LEB†

* Syllables used for the serial learning.

† Contained in published lists of low associative value.

The syllables in the first column of Table I served a dual purpose. They were learned by all Ss in the serial learning and were the stimulus-items for the subsequent paired-associates for experimental Groups A and D, as described below. The syllables in the second column were chosen as being phonetically similar to those in the serial list and were learned as stimulus-members in the paired-associates learning by Groups B and E. The third column of Table I contains the syllables constructed to be physically similar in appearance to those in the serial list and were learned as the paired-associates stimulus-items for Groups C and F. For each trio of syllables on the same row in the first three columns of Table I, the same paired-associate response-member was used, and these are contained in the last column of Table I. The fourth and fifth columns of this table indicate the nature of the change in the syllable of the first column in order to produce the forms shown in the second and third columns used as homophones and as physically similar stimuli.

⁶ E. R. Hilgard, Methods and procedures in the study of learning, Ch. 15 in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951.

⁷ C. I. Hovland, Experimental studies in rote-learning theory. I. Reminiscence following learning by massed and distributed practice, *J. Exper. Psychol.*, 22, 1938, 201-224.

Eight different orders of the eight syllables were used in the serial learning, and the *Ss* were randomized with regard to the order group into which they were assigned. Each order group was equally represented in each of the six experimental groups. These experimental groups for the paired-associates learning were arranged as follows, with eight *Ss* in each group.

In each of the experimental groups there was one *S* from each of the order groups as described above. In Conditions A, B, and C, the paired-associates learning followed the serial learning by approximately 4 sec. In Conditions D, E, and F, a 24-hr. interval separated the serial from the paired-associates learning. In learning the pairs each *S* was presented as stimuli the items he had previously learned in Positions 2, 3, 5, 6, and 8 of the serial list or variants of these stimuli according to whether the *S* was in the 'identical, homophone, or physically similar' group. It was judged that five pairs of items in paired-associates learning would be optimal for a group of *Ss* of the age group used in this experiment, and the particular serial positions

TABLE II
EXPERIMENTAL CONDITIONS

Stimulus-form	Time after serial learning	
	immediate	24 hr. interval
identical	A	D
phonetically similar (homophones)	B	E
physically similar in appearance	C	F

of the syllables chosen allowed sufficient data to be obtained regarding position effects to test the experimental hypotheses adequately.

Procedure. The *Ss* were dealt with individually and practice periods were given in learning letter-number combinations. Detailed but routine instructions were given regarding the procedure for serial learning and for paired-associates learning. Because of the ages of the *Ss*, considerable attention was given to the instructions and to ascertaining that all *Ss* understood the procedure exactly. The experiment was conducted as a challenging game for the *Ss*, and in the final stage of paired-associates learning the *Ss* were acting to "try to beat the flap" that opened to expose the response member of each pair.

The syllables of Column 1 in Table I were presented as the list for serial learning and in different orders for the various *Ss*. These eight orders were so arranged that each syllable occupied each position only once and followed each of the other syllables only once. The serial list of eight nonsense syllables was presented at the constant rate of 2 sec. per syllable with a delay of 6 sec. between the terminal syllable and the cue syllable. Incomplete learning was desired and so the serial learning was terminated for each *S* when he had given five correct anticipatory responses from the possible eight responses in the list. Records were kept of trials and syllable errors by an independent observer.

The paired-associates learning began immediately for three groups and was delayed 24 hr. for three groups. The stimulus-pairs for any one *S* were determined by the serial group and the experimental group to which he had been randomly assigned. The pairs to be learned had as their stimulus-members the five syllables learned in

Positions 2, 3, 5, 6, and 8 or the variants of them according to the experimental design above. For all Ss the respective five pairs of syllables were presented in the apparatus at the rate of 3 sec. for the stimulus-member and 3 sec. for both the stimulus- and response-members together. The order of presentation of the pairs varied from trial to trial. The method of adjusted learning was used for the paired-associates, and each syllable card was removed from the presentation system when the response had been correctly anticipated three successive times. Complete records of trials, correct responses, and errors were made by an independent E. An error was taken as an omission or any incorrect substitution. Errors were used for final measures instead of correct responses or correct responses-minus-errors.

RESULTS

The errors in the serial learning and paired-associates learning were separately analyzed. Analysis of the errors for each of the eight positions in serial learning showed that positional differences were significant and that the mean total error scores for the eight positions are described by the bow-shaped function found in previous studies.⁸ Critical to our investigation, however, is the persistence of the bow-shaped error-function through the criterion trial of serial learning. Error responses on the last trial were therefore subjected to a separate analysis. Table III contains the

TABLE III
MEANS AND STANDARD DEVIATIONS OF TOTAL ERRORS FOR EACH SYLLABLE
POSITION ON THE CRITERION TRIAL

	Position in the serial list							
	1	2	3	4	5	6	7	8
mean	.06	.17	.29	.64	.67	.60	.23	.00
SD	.08	.16	.81	1.56	1.58	1.47	.65	.00
rank order	2	3	5	7	8	6	4	1

summary data regarding errors on the last trial. Complete analysis of variance was made for the last trial errors. Position effect was found to be significant at the 1-% level. The tables for these data are not presented here because of their complexity of detail and because the particular analyses are not of primary importance in the study. The results substantiate the presence of differences in syllable position learning immediately prior to the presentation of the paired-associates.

The six experimental conditions of the paired-associates learning permitted the investigation of the two primary experimental variables: (1) two conditions of time-delay between the learning sessions; and (2) three

⁸ C. I. Hovland, *ibid.*; Hull et al., *op. cit.*; L. B. Ward, Reminiscence and rote learning, *Psychol. Monog.*, 49, 1937 (No. 220), 1-64

conditions of the form of the paired-associates stimulus-member relative to the serial learning: identical, homophones, or physically similar in appearance. In analyses of such data the problem becomes one of choosing a sufficiently sensitive dependent variable. Not all of the measures of learning efficiency available on the paired-associates items would be sensitive enough to reflect transfer effects because of the rapid decay of the inhibitory potential developed during the serial learning. Complete analysis of the total errors to criterion did not indicate a significant position effect in all groups, probably because any transfer or generalization of negative potential did not extend through the period of learning the five paired-associates to criterion. In this analysis of total errors the time delay factor was statistically significant at the 5-% level.

Because of this contamination in the measures of total errors to criterion, the dependent variable of errors within the first four trials of the paired-associates learning was used. This measure was based on the assumption that the inhibitory potential, if present, should be at its maximal strength immediately following the serial learning. This consideration was indicated by Hull's 13th postulate which describes inhibitory potential as dissipating rapidly in time.⁹ Also Ward's studies on reminiscence indicate that the maximal dissipation of the inhibitory potential occurs within a period of 2 min. following the terminal serial learning trial.¹⁰ Under the experimental condition of immediate paired-associates learning in this study, the first four trials of the five pairs of nonsense syllable were completed within approximately 142 sec. after the last trial of serial learning. This feature is the justification for analyzing the first four trials of paired-associates learning, *i.e.* they were completed within a time-period approximately the same as that found by Ward for dissipation of inhibition. Actually, a separate analysis was carried out for the first trial of learning by paired-associates with results similar to those found for the first four trials.

The means and *SDs* of the error responses during the first four trials by paired-associates for the six experimental groups are shown in Table IV according to the position occupied by the stimulus-member in the original serial learning. Each mean entry in Table IV, is based on an *N* of eight *Ss* with each *S* having learned a different pair of nonsense syllables. The five position mean values for each experimental condition are shown graphically in Fig. 1. Inspection of the graphs in Fig. 1 reveals that three

⁹ Hull, et al., *op. cit.*, 67.

¹⁰ Ward, *op. cit.*, 63 f.

curves, for Conditions A, B, and D, have bow-shapes similar to those found for serial position effects. The curve for Condition A is the most similar

TABLE IV

MEANS AND SDs OF ERRORS OBTAINED DURING FIRST FOUR TRIALS FOR THE SIX PAIRED-ASSOCIATES CONDITIONS

			Position in serial list occupied by identical or similar syllable as that of stimulus-member				
Condition			2	3	5	6	8
Immediate	A identical	M	2.12	3.12	3.50	3.25	2.75
		SD	2.21	1.70	2.00	2.35	1.87
	B homophones	M	2.62	3.37	3.62	3.75	3.00
		SD	1.97	1.97	1.97	1.23	2.00
	C phys. sim.	M	3.50	3.62	3.37	2.87	3.25
		SD	2.45	1.37	2.42	2.98	1.87
24-hr. delay	D identical	M	3.50	3.62	3.87	3.50	3.50
		SD	2.00	1.32	.94	2.45	2.83
	E homophones	M	3.87	3.62	3.50	4.00	3.87
		SD	.94	1.37	2.45	.00	.94
	F phys. sim.	M	3.87	3.00	3.62	3.50	3.37
		SD	.94	3.74	1.97	2.00	2.42

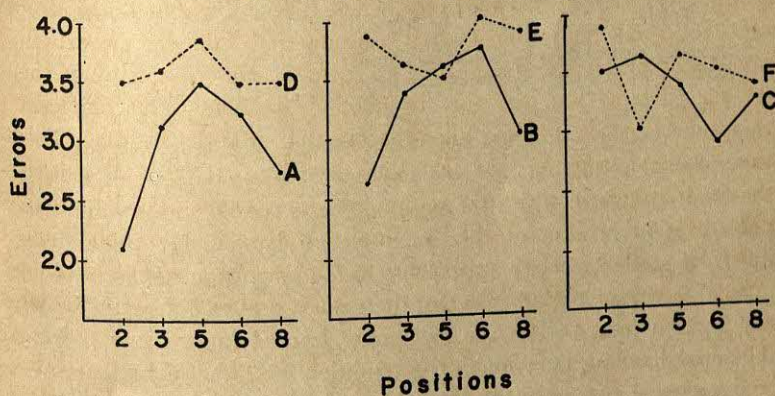


FIG. 1. MEAN ERROR RESPONSES OBTAINED DURING FIRST 4 TRIALS OF PAIRED-ASSOCIATES LEARNING FOR EACH EXPERIMENTAL GROUP

The first graph is for the identical condition, second for homophones, and third for physically similar stimuli. The heavy lines are for the immediately successive learning, and the dotted lines are for 24-hr. delay.

to the expected function in that its maximum is for Position 5, while the maximum for Condition B is found at Position 6.

The analysis of variance was used to test the significance of the data shown in summary form in Table IV. Bartlett's test gave no reason to

reject the hypothesis of homogeneity of variance. The full analysis of the data inappropriately pools the radically different conditions of the time-delay with measures of position effect and serves to mask the more immediate effects of position differences within the condition of immediately successive learning. This general analysis was performed, however, for the data on the first four trials of the six experimental conditions for the paired associates. This analysis is shown in Table V, where it is seen that time-delay effects and position-time interaction were statistically significant. In the light of our hypothesis that the positional effects, indicating the

TABLE V
ANALYSIS OF VARIANCE FOR ERRORS DURING FIRST FOUR TRIALS OF THE
SIX EXPERIMENTAL CONDITIONS

Source of variation	df	Mean square	F
position	4	.88	1.40
time	1	11.27	19.10†
form	2	1.25	2.12
position×time	4	1.74	2.95*
position×form	8	.71	1.20
time×form	2	1.30	2.20
position×time×form	8	1.02	1.73
within subgroups (error)	210	.59	
Total	239		

* Significant beyond the 5-% level of confidence.

† Significant beyond the 1-% level of confidence.

carry-over of the inhibitory potentials, would vary as a function of the experimental conditions, a non-significant *F*-ratio would be expected for the total analysis in which the position effects for all six conditions were pooled. Substantiation of this interpretation is available from the significantly greater variance attributable to the interaction of time and position. This interaction indicates that there was a greater differential between the position means when the pairs were learned immediately following the serial learning than under the condition of 24-hr. delay, a result to be expected according to the general hypotheses under investigation.

To explore further the implications of the significant interaction of time and position, an analysis of variance was conducted for each condition individually. Position was the independent variable tested for each condition. Based upon 4 degrees of freedom, the position mean squares were compared with the within-group variance for each condition. The relevant data of the six analyses are presented in Table VI. In this more detailed and more appropriate analysis significant *F*-ratios were obtained

for the variances attributable to syllable position differences for Conditions A and B, both well beyond the 5-% level of confidence. The two bow-shaped curves of the position means expressed in terms of error-responses obtained during the first two minutes of the paired-associates learning (Figure 1) may therefore be considered as indicating significant trends. The trends shown in the curves for Conditions A and B are in agreement

TABLE VI

RATIOS OF POSITION-VARIANCE WITHIN GROUP-VARIANCE, DEGREES OF FREEDOM, AND *F*-RATIOS FOR THE ERROR RESPONSES OBTAINED DURING FIRST FOUR TRIALS FOR ALL EXPERIMENTAL CONDITIONS

	Condition	Variances	df	<i>F</i>
Intermediate	A identical	2.29/.59	4/35	3.88†
	B homophones	1.71/.49	4/35	3.49*
	C phys. similar	.66/.75	4/35	—
24-hr. delay	D identical	.21/.59	4/35	—
	E homophones	.34/.27	4/35	1.26
	F phys. similar	.84/.82	4/35	1.02

* Significant beyond the 5% level.

† Significant beyond the 1% level.

with our prediction based on the presumed transfer and generalization of the inhibitory potentials generated in the serial learning.

As a further way of analyzing the data and as a more extreme test for the relative significance of the experimental factors introduced, the experimental condition of physically similar syllables was eliminated and the remaining sources of variation tested. These results are shown in Table VII. This procedure of eliminating experimental conditions for clearer demonstration of significant results is not recommended as a generally valid procedure, and it is to be used with caution. Table VII, however, again indicates the significant effect of the serial position in which the stimulus-syllables had been originally learned, and it also demonstrates the fact that the effects generalize to homophones of the original syllables.

DISCUSSION

Though the data and results obtained here may be interpreted in terms of positional-difficulty in rote serial learning, we believe that an interpretation in terms of an inhibitory potential is the more theoretically sound. It is from this position that much of the currently available experimental data and extensive theory have been developed.

The general formulation of the theory involving inhibition of delay as a mechanism underlying the serial-position phenomena is attributed to Hull and his associates.¹¹ We have postulated additional characteristics of the hypothetical inhibitory potential to the effect that, within limits, it will transfer to new situations and will generalize to stimulus-objects on a gradient of similarity. The evidence appears to confirm these postulates.

The phenomena of transfer and generalization of the inhibitory potential appear to be important in terms of theoretical interpretations of learning behavior. Under the conditions of this experiment, however, the evidence available indicates that while these characteristics of the inhibitory potential

TABLE VII
ANALYSIS OF VARIANCE OF ERRORS DURING FIRST FOUR TRIALS OF PAIRED-
ASSOCIATES LEARNING UNDER CONDITIONS A, B, D, AND E

Source of variation	df	Mean squares	F
position	4	2.00	4.12†
time	1	13.22	26.98†
form	1	2.50	5.10*
position×time	4	1.90	3.88†
position×form	4	.51	1.04
time×form	1	.23	—
position×time×form	4	.11	—
error	140	.49	—
Total	159		

* Significant, 5% level.

† Significant, 1% level.

are definite they are short-lived. This is evidenced by the lack of any differential effects after a period greater than 2 min. Because of this temporal limitation, these characteristics of transfer and generalization may not be useful in general descriptions of negative transfer and interference.

The later formulations by Hull regarding the mechanism of inhibition emphasize the attachment of the negative potential to the response member in associative learning.¹² It is obvious from this formulation and from the results of the present experiment that a further investigation is needed to determine the extent of transfer and generalization of inhibitory potential associated with response members in paired-associates learning. From the experimental results obtained by Underwood it is judged that significant effects may well occur in transferring and generalizing inhibitory

¹¹ Hull, et al., *op. cit.*

¹² Hull, *Principles of Behavior*, 1943.

potential from serial learning to the response members in paired-associates learning.¹³

SUMMARY AND CONCLUSIONS

Forty-eight school children learned a serial list of eight nonsense syllables to a criterion of partial learning. Following the serial learning the Ss learned sets of five paired-associates. The stimulus-members of the paired-associates were chosen from the 2nd, 3rd, 5th, 6th, and 8th positions in the original serial list. The Ss were divided into six experimental groups with respect to the factor of similarity of the new stimulus to items in the serial list and with respect to the delay in time before beginning the paired-associates learning. The following relationships obtained for the six groups: identical-immediate; identical-delayed; phonetically similar-immediate; phonetically similar-delayed; physically similar-immediate; and physically similar-delayed.

Within the first four trials of the immediately following paired-associates learning, significant effects of transfer for the identical syllables and significant generalization to the homophones were demonstrated. Within the framework of this experiment, the following conclusions are presented.

- (1) Inhibitory potential developed in rote serial learning transfers to new learning situations.
- (2) Such inhibitory potential generalizes in verbal learning on the basis of the homophony of the material.
- (3) The transfer and generalization of the inhibitory potential are short-lived and do not occur after approximately 2 min. following the original learning.
- (4) Inhibitory potentials thus appear to have significant characteristics beyond those originally postulated in learning theory.

¹³ B. J. Underwood, Associative transfer in verbal learning as a function of response similarity and degree of first-list learning, *J. Exper. Psychol.*, 42, 1951, 44-53.

THE INFLUENCE OF PHOTOMETRIC BRIGHTNESS ON JUDGMENTS OF SIZE

By EDWARD J. ROBINSON, Boston University

Interest in the effect of brightness upon judgments of size antedates the science of psychology. While many writers have stated that brighter objects tend to appear larger, no systematic relationships are reported in their studies.¹ The two studies which have direct bearing on the problem are inconclusive and contradictory.² Holway and Boring found that their largest brightness difference produced approximately 25% error in equating a variable with a standard light source. Pheiffer, on the other hand, found his largest brightness difference produced approximately 3% error. In addition, he noted that the brighter object did not appear larger for all the standard stimulus-sizes used.

Two experiments are reported here. The first is concerned with the effect of photometric brightness on judgments of size under binocular and monocular conditions. The second experiment was designed to determine whether the relationship in the first experiment obtains for various absolute levels of photometric brightness.³

METHOD

Apparatus. The apparatus is pictured in Fig. 1. Basically it consisted of two

* Accepted for publication August 5, 1953. This study is based on a dissertation submitted to the Graduate School of Boston University in partial fulfillment of the requirements of the Doctor of Philosophy degree. It was conducted in the Psychology Section of Boston University, Optical Research Laboratory, under contract with the Photographic Reconnaissance Laboratory, Wright Air Development Center, USAF (Contract AF33-038-15615). The research was directed by Dr. L. J. Reyna.

¹ H. von Helmholtz, *Physiological Optics*, vol. 2, J. P. C. Southall, (ed.), 1924; N. R. F. Maier, The illusion of size in pin-hole vision, this JOURNAL, 41, 1929, 291-295; C. E. Ferree, and G. Rand, Size of objects in relation to their visibility and to the rating of vision, *Arch. Ophthalm.*, 4, 1930, 37-72; T. Amano and H. Shigeno, Preliminary study on the perception of visual size, *Acta Psychol., Keijo*, 2, 1933, 1-12, *Psychol. Abst.*, 7, 1933, 3135; Curt Berger, A comparison under different degrees of brightness, of minimum visual distances of two luminous points and of a broken circle, this JOURNAL, 55, 1942, 354-370; G. L. Walls, Factors in human visual resolution, *J. Opt. Soc. Amer.*, 33, 1943, 487-505.

² A. H. Holway and E. G. Boring, The dependence of apparent visual size upon illumination, this JOURNAL, 53, 1940, 587-589; C. H. Pheiffer, A psychophysical investigation of the discriminial limits of size-differences of discs of light, *Ohio State University Abstracts of Doctoral Dissertations*, 1948-1949, No. 60, 243-247.

³ The term 'photometric brightness' is used in this paper in accordance with the usage of the American Standards Association. It will refer to the stimulus and not

stimulus-lights and *O*'s station.⁴ *O*'s head was held firmly in position by a modified welder's mask. A hinged opaque window in the mask, which could be opened and closed by *O*, permitted him to view the stimulus-lights at will. On the left and just below the mask there was an intercommunication unit through which *O* could

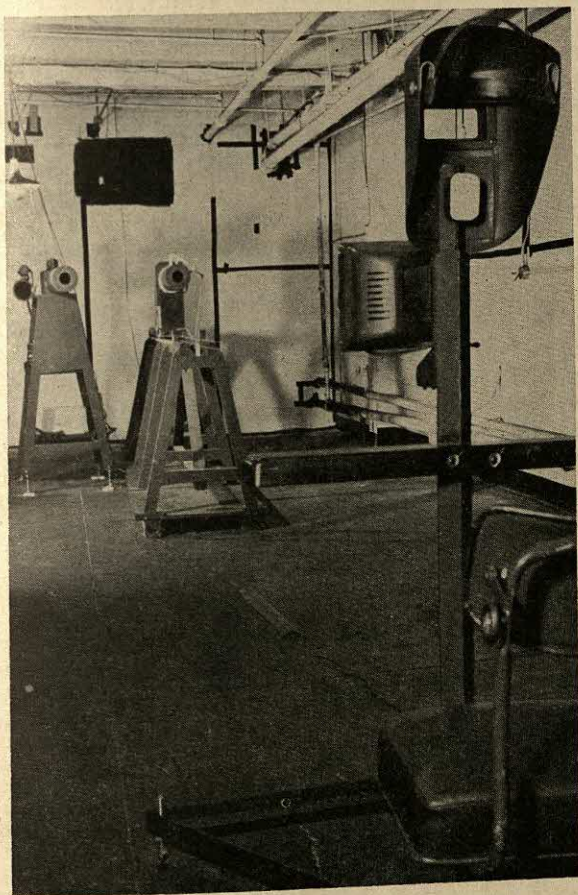


FIG. 1. GENERAL LAYOUT OF THE APPARATUS

speak to *E* who was located behind the two lights. The apparatus was in a light-proof room.

Both stimulus-lights were evenly illuminated disks, spatially fixed, and located at a distance of 25 ft. from *O*. They were at a fixed distance to exclude accommoda-

to its appearance to the subject. The term 'brightness' will be used as a contraction of 'photometric brightness.'

⁴The apparatus was designed by John Coules and E. J. Robinson for a program of research in visual discrimination.

tion as a variable. The light on *O*'s left, fixed in size, was the standard (*SL*); that on his right, variable in size, was the variable (*VL*). Each light had a wooden housing and a system of projecting tubular baffles. The tubular baffles were used to eliminate stray light from the test-room and to protect *O* from unwanted cues. Each baffle tube, 25 in. in length and 6 in. in diameter, was made of black paper bakelite. Five circular aluminum baffles, with 2.9-in. diameter opening in the center of each, were placed at 5-in. intervals throughout the length of each tube. The entire apparatus was painted a flat black inside and out.

The lamps (6-v., 3.75 amp.) within each housing unit were mounted on a rotatable supporting base because the filaments of the lamps produced a dark spot on the opal disks. Thus the lamps could be rotated through 360° and homogeneous illumination of the disk be attained.

The illumination on the opal glass could be varied by moving the lamp in the housing unit. A calibrated rod connected to the base plate permitted *E* to obtain the desired value of illumination. The change in illumination followed the inverse square law. In this way, a brightness gradient was obtained without any color temperature change. In addition, provision was made for Wratten neutral-density filters to be placed over the first baffle on the front of both the standard and variable. These filters served two purposes: (1) they permitted *E* to obtain the brightness levels desired; and (2) they reduced the scattering of the light and eliminated the 'halo' effect around the lights.

An aluminum knife-edge was installed between the housing and the opal glass disk to eliminate light scattering. The knife-edge in the standard was 2 in. in diameter; in the variable it was 3 in. This difference was necessary because the diaphragm in front of the opal glass in the variable was at times opened wider than the 2-in. The diaphragm system was calibrated from 1.600 in. to 2.400 in. in 0.025-in. increments.

With maximal diaphragm opening and minimal distance between the light bulb and the opal glass in the variable, a non-uniform distribution of illumination on the opal glass occurred. To correct this, a white, Bristolboard cylinder, 2 in. in length and 3 3/8 in. in diameter, was installed at the entrance to the baffle tube on the inside wall of the housing of the variable unit.

The electrical system (100-v. AC) consisted of a voltage regulator, a variable transformer (Variac), and a 6-amp. filament transformer. All three were in series. The two lamps were in parallel from the last unit in the series.

PROBLEMS

Experiment 1. The purpose of Experiment 1 was to determine the functional relationship between brightness and judgments of size under binocular and monocular conditions.

Procedure. Preliminary investigation indicated (1) a 0.025-in. interval was necessary to insure differential responses in size; and (2) brightness-ratios of 7.5:1 and less had to be used to obtain differential responses of larger and smaller. The ratios used in this experiment were: 1:6.25, 1:2.5, 1:1, 2.5:1, and 6.25:1.

In this experiment only one brightness-level was used (13.0 foot-lamberts). In

each stimulus-combination one of the light sources had a brightness of 13.0 foot-lamberts and the other light source was decreased to yield one of the five brightness-ratios. All of the data were plotted in terms of the 'smaller' responses to the VL.

Two women, undergraduates in psychology, with visual acuity of 20/20 in each eye, served as *O*. They participated in approximately 20 daily sessions spaced over a period of 8 weeks. The instructions required *O* to report whether the right or left light was the smaller. The method of constant stimulus-differences was used to present the 85 stimulus-combinations at all of the 5 brightness-ratios. *O* was run first under binocular conditions (30 presentations per stimulus-combination) and then under monocular conditions (20 presentations per stimulus-combination). Under the monocular conditions *O* used her dominant eye. Both *O*s had right-eye dominance.⁶

Experiment 2. Experiment 2 was designed to determine how the size-brightness relationship is affected when the absolute level of brightness is changed.

Procedure. In this experiment three levels of brightness were used and were obtained by placing filters at the end of the baffle tubes.⁶ The brightness ratios, apparatus, subjects, and experimental room were the same as in Experiment 1. A modified method of limits was used to obtain the various points of subjective equality.⁷ *O* was instructed to state whether the light on the right was smaller or larger than the one on the left.

The stimulus-light, set at a particular brightness-ratio, was first presented *O* binocularly at 13.0 foot-lamberts and at 5 ratios in the order: 6.25:1, 1:1, 1:2.5, 1:6.25, and 2.5:1. The same order of ratios was then repeated from monocular observation. This procedure was then repeated for 3.33 and 0.035 foot-lamberts.

RESULTS

Experiment 1. The effects of brightness on judgments of size under binocular conditions are shown in Fig. 2.⁸ With the 1:1 ratio as a reference curve, it can be seen that the 6.25:1 and 2.5:1 curves are displaced to the right, indicating that although the variable stimulus-light was greater than 2.000 in. it was still being responded to as smaller. Conversely, the 1:6.25 and 1:2.5 curves are displaced to the left of the reference curve.

⁶ The first 15 presentations were plotted and compared with the last 15 presentations for both *O*s under binocular conditions to determine whether there were any systematic changes between responses over the course of the experiment. Likewise, the first 10 presentations were compared with the last 10 presentations for both *O*s under monocular conditions. Inspection of these curves revealed no major trend differences.

⁷ Increasing the absolute level of brightness above 13.0 foot-lamberts brought with it an undesirable 'halo' effect around the light sources. It was decided for the present study to decrease only the brightness levels.

⁸ Dr. G. A. Miller, personal communication, 1953.

⁸ The results for the other *O* were similar. To conserve space they are not presented.

indicating, although the variable light was smaller than the standard, it was still being responded to as larger than 2.000 in.⁹

Fig. 3 shows the curves for the same *O* under monocular conditions. The same relationship noted above is evident.

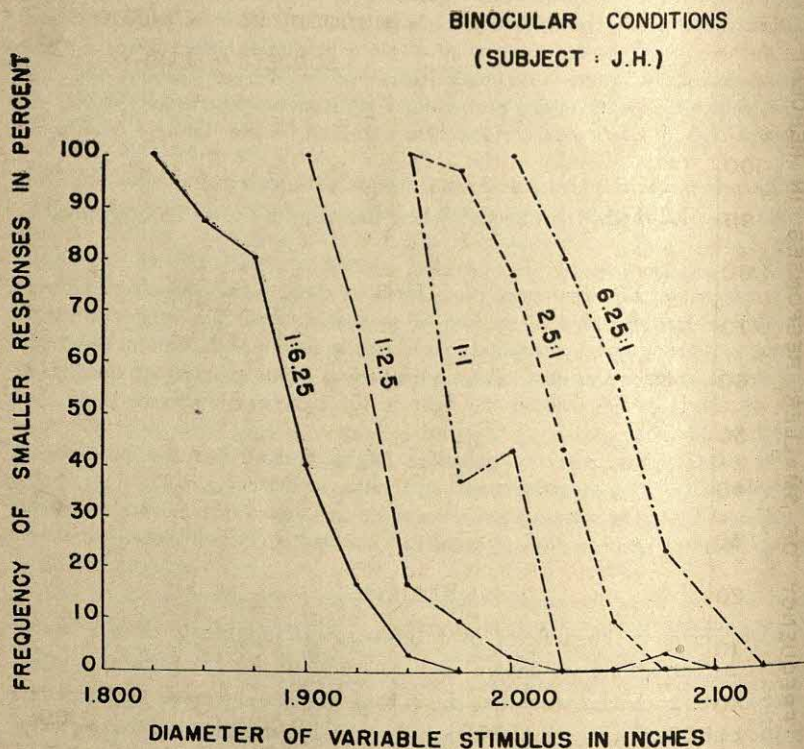


FIG. 2. PERCENTAGE FREQUENCY OF 'SMALLER' RESPONSES AS A FUNCTION OF SIZE OF THE VARIABLE STIMULUS-LIGHT AND BRIGHTNESS-RATIOS
(Every point based on 30 responses)

To show the effect of brightness on judgments of size more clearly additional curves were derived from Fig. 2 and 3. In these curves the size of the variable was held constant at 2.000 in. and the percentage 'smaller' responses was plotted against the logarithm of the brightness-

⁹ Based on the assumption that the above curves are linear functions over the range of sizes and brightness ratios employed in this study, the following relationship was found to fit the data: Percentage 'smaller' responses = $2535 \times 107 \log (\text{brightness ratio}) - 1250 \text{ diameter}$.

ratios. Fig. 4 shows that as the relative brightness of the variable is decreased, the percentage 'smaller' responses to it gradually increases.

In addition to the determination of the usual *S-R* type of relationship, *S-S* functions are necessary to ascertain how one stimulus-condition varies

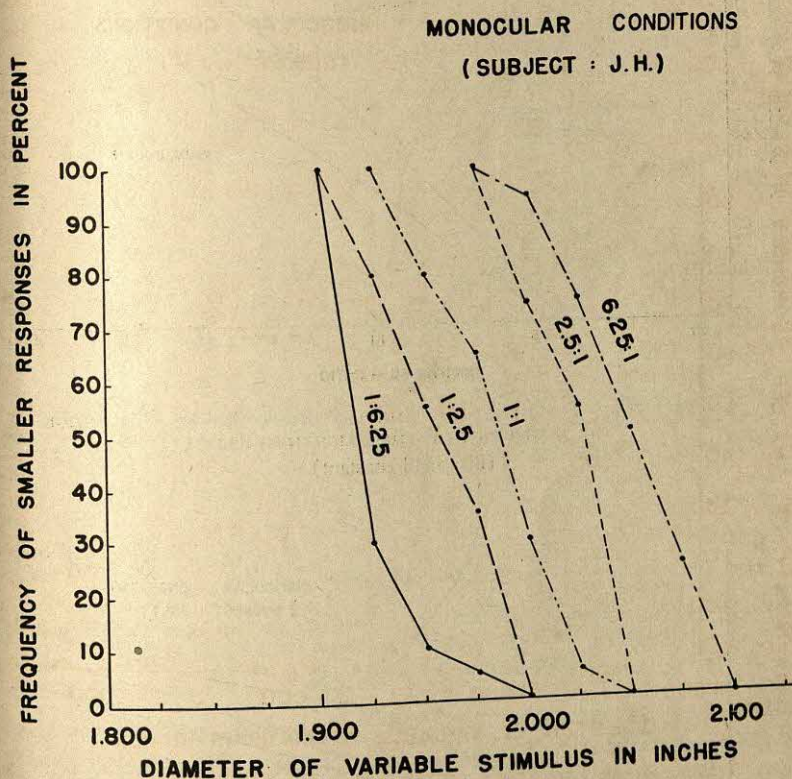


FIG. 3. PERCENTAGE FREQUENCY OF 'SMALLER' RESPONSES AS A FUNCTION OF SIZE OF THE VARIABLE STIMULUS-LIGHT AND BRIGHTNESS-RATIOS
(Every point based on 20 responses)

as a function of another stimulus-condition in producing a constant response.¹⁰ To obtain this type of relationship the 50% (and the 25% and 75%) 'smaller' response-frequencies were taken for each brightness-ratio and plotted against the corresponding size of the variable. Fig. 5 was derived in this manner from Fig. 2.¹¹

¹⁰ C. H. Graham, Behavior, perception, and the psychophysical methods, *Psychol. Rev.*, 57, 1950, 108-120.

¹¹ Results for *JH* under monocular conditions are very similar.

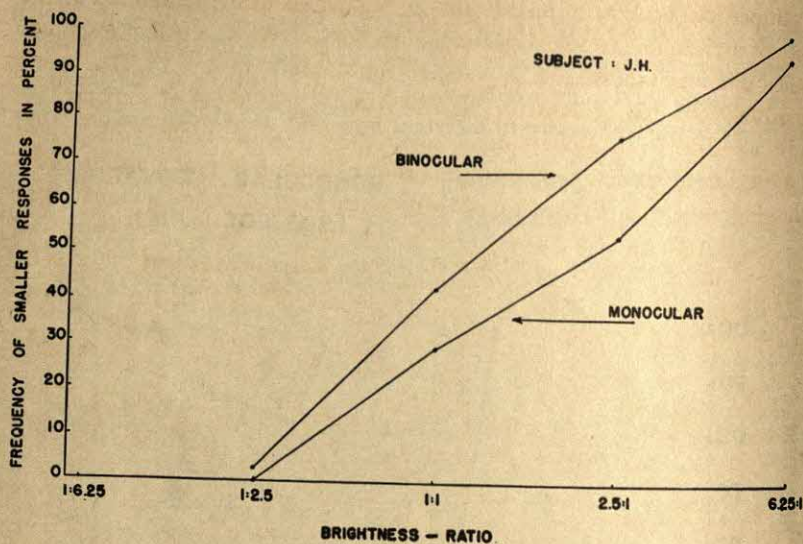


FIG. 4. PERCENTAGE FREQUENCY OF 'SMALLER' RESPONSES AS A FUNCTION OF THE LOGARITHM OF THE BRIGHTNESS-RATIOS
(Size held constant)

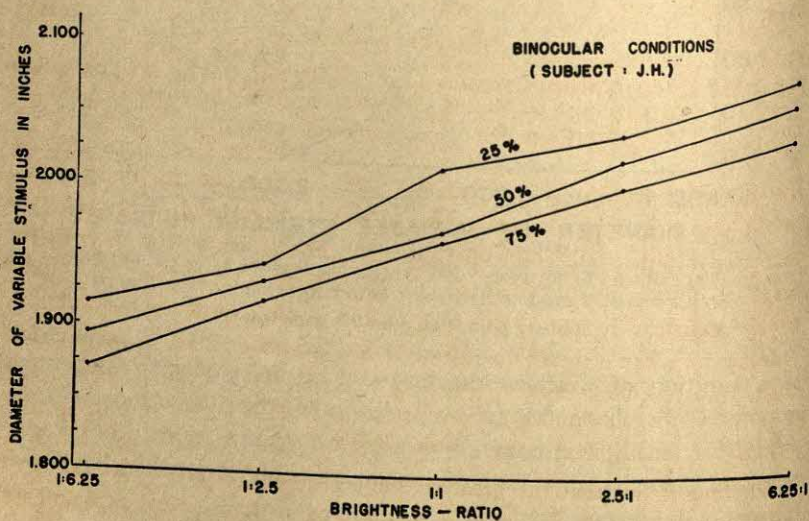


FIG. 5. SIZE OF THE VARIABLE STIMULUS-LIGHT AS A FUNCTION OF THE LOGARITHM OF THE BRIGHTNESS-RATIOS
(Response held constant)

Inspection of Fig. 5 reveals that a regular, consistent decrease in the size of the variable is necessary to obtain a 50-% level of 'smaller' responses when the brightness-ratio is varied from 6.25:1 to 1:6.25. The same consistent trend is evident for the 25-% and 75-% levels of response.

Experiment 2. The effect of decreasing the absolute level of brightness on the size-brightness relationship is shown for one *O* (*JH*) in Fig. 6. As this Figure shows, the relationship between brightness and judgments of

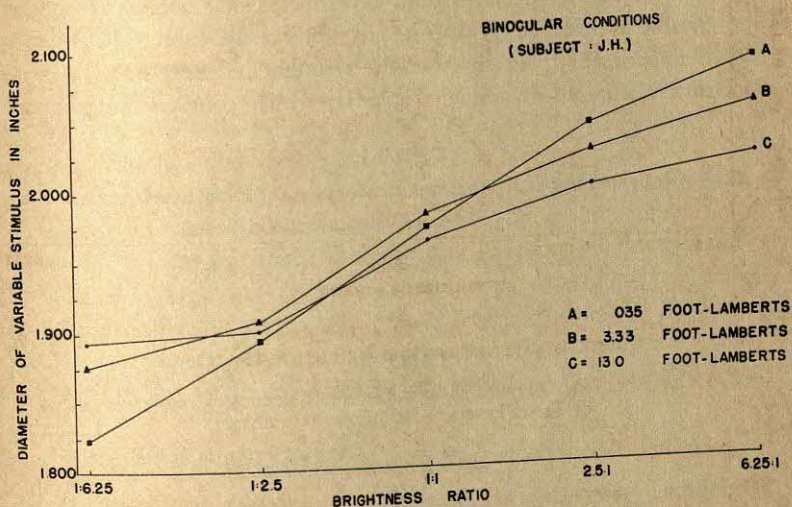


FIG. 6. SIZE OF THE VARIABLE STIMULUS-LIGHT AS A FUNCTION OF THE LOGARITHM OF THE BRIGHTNESS-RATIOS AT THREE LEVELS OF BRIGHTNESS
(Every point based on 60 responses)

size noted in Experiment 1 remains essentially unchanged. Examination of the curves in Fig. 6 suggests that as the absolute level of brightness is increased (in this case from 0.035 to 13.0 foot-lamberts) the effect of the increased brightness-ratio is decreased. A more precise statement regarding this relationship would require the methodology employed in Experiment 1 at absolute levels greater than, as well as less than, 13.0 foot-lamberts.

DISCUSSION

Pheiffer found that the greatest change brought about by the influence of brightness was on the order of $\frac{2}{32}$ in., or in terms of his experimental conditions, about a 3% change. The results of this experiment agree very closely with this figure in that the greatest effect due to brightness was

found to be on the order of 3-4%. It is not clear what the relevant variables are that could account for the discrepancy between these results and those of Holway and Boring. One possible factor might be that in the Holway and Boring experiment *O*s had to rotate their heads 90° to compare the two stimulus-lights. It is possible that the change in pupil size brought about by this procedure is responsible for the magnitude of error which they report.

It is evident in the results of both experiments that our *O*s responded to the variable as being larger than its physical size. In Table I the constant errors in Experiment I, which theoretically might be expected to be zero, are consistently minus in value. Table II, which contains the constant

TABLE I
CONSTANT ERRORS UNDER BINOCULAR AND MONOCULAR CONDITIONS (1:1 RATIO)

O	Constant error (in inches)	
	Binocular	Monocular
JH	-.032	-.013
CF	-.028	-.038

TABLE II
CONSTANT ERRORS UNDER MONOCULAR AND BINOCULAR CONDITIONS
OF THREE LEVELS OF BRIGHTNESS (1:1 RATIO)

Conditions	O	Constant errors (in inches)		
		13.0 foot-lamberts	3.33 foot-lamberts	0.035 foot-lamberts
Binocular	JH	-.038	-.020	-.027
	CF	-.051	-.051	-.069
Monocular	JH	-.034	-.019	-.025
	CF	-.029	-.055	-.085

errors for both *O*s in Experiment 2, shows the same bias. The magnitude of the bias in the latter is even greater in many instances.

A possible reason for the increase in bias in Experiment II is the change in instructions. In this experiment it will be recalled *O* was required to state whether the light on her right was larger or smaller than the one on her left. The light on the left (the standard), although it was not explicitly stated, became the standard. Under similar conditions many other investigators have found the stimulus to be judged tends to be responded to as larger than the standard stimulus.¹² This phenomenon is generally referred to as 'space error.'

It does not appear feasible, however, to attribute this bias to space

¹² Holway and Boring, *op. cit.*, 587-589; Pheiffer, *op. cit.*, 243-247.

error in Experiment 1. The instructions there required *O* to report which light appeared 'smaller,' the one on her left or the one on her right. Under these conditions neither light was designated as the standard or comparison stimulus. This was verified after the experiment when the *O*s reported that they thought both lights were being varied in size as well as brightness. In addition, they evidenced surprise when they found that the light on their left (the standard) had never changed in size.

One suggestion to account for this bias is derived from an analysis of the data under monocular conditions. In Tables I and II the direction of the bias can be seen to be the same for both *O*s. These *O*s were required to use their preferred eye under monocular conditions, and for both *O*s this was their right eye. It is significant that the position of the variable corresponded to the dominant eye of both *O*s.

Additional support for the notion that eye-dominance might have been the relevant variable for the bias is found in an experiment by Coules, who used the same apparatus to investigate distance judgments and their relationship to brightness.¹³ He also found this systematic bias; namely, the *O* using his right eye as the preferred eye responded to the variable stimulus as being nearer than it actually was. On the other hand, the *O* using his left eye as the preferred eye responded to the standard as being nearer.

SUMMARY AND CONCLUSIONS

Two experiments are reported. In Experiment 1 the effect of brightness on judgments of size under binocular and monocular conditions was investigated. In Experiment 2 the purpose was to determine whether the relationship found in Experiment 1 held for different absolute levels of brightness.

The apparatus permitted the presentation of two homogeneous disks of light: a standard stimulus-light of 2.000-in. diameter and a variable that could be changed from 1.600 to 2.400 in. in diameter. Particular attention was paid to control of the stimuli likely to influence size judgments. The factors controlled included: accommodation, stray light, homogeneous illumination of the stimulus-disk, light scatter, and color temperature change of the light sources.

The stimuli were presented according to the method of constant stimulus in Experiment 2, and a modified method of limits in Experiment 2. S-R relationships were obtained between brightness and the percentage of

¹³ J. Coules, The influence of photometric brightness on judgments of distance. Unpublished Ph.D. dissertation, Boston University, 1953.

'smaller' responses. The curves indicated that brightness operated systematically to make the brighter object appear larger. This relationship was evident under binocular and monocular conditions. In addition, S-S functions were derived and these curves indicated that a regular, consistent decrease in the size of the variable was necessary to obtain a 50-% level of 'smaller' responses as the brightness-ratio varied from 6.25:1 to 1:6.25.

Under the conditions of the experiment (2.000 in. standard) the increase in size due to brightness was on the order of 3-4%. Change in pupil size was suggested as the variable accounting for the difference between the present results and those of Holway and Boring.

In Experiment 2, where the size-brightness relationship was investigated under different levels of brightness, it appears that as the absolute level of brightness is increased (0.035 to 13.0 foot-lamberts) the effect of the brighter object in any particular brightness-ratio is decreased.

THE EFFECT OF ANGLE OF TILT UPON THE TRIGONOMETRIC RELATIONSHIP OF PRECISION AND ANGLE OF LINEAR PURSUIT-MOVEMENTS

By GEORGE E. BRIGGS, RICHARD F. THOMPSON, and
W. J. BROGDEN, University of Wisconsin

This paper presents the results of an experiment on the effect of angle of tilt, *i.e.* deviation of the tracking surface from the horizontal, on the relation of precision and angle of linear pursuit-movements. Corrigan and Brogden found this latter relation to be described adequately by the following trigonometric equation:¹ $y = a + b \cos 2x + c \sin 2x$, in which y = precision of right-arm movements in terms of group mean frequency of stylus contact, x = angle from the body at which the movement is made,² a = the constant that determines the baseline of the curve (mean frequency of stylus contact for all angles), and b and c are the constants by means of which the amplitude (d) and the phase angle ($2e$) of the curve are determined [$d = (b^2 + c^2)^{1/2}$; $\cos 2e = c/d$.] For this and other studies, the tracking surface was maintained in the horizontal plane or at 0° tilt. An initial experiment on the effect of tilt, and angle from the body on the precision of linear pursuit movements revealed both variables to be significant determiners of precision in tracking. The Greco-Latin square design of this experiment, however, did not provide adequate measurement of the effect of the two experimental variables. In the first place, a significant interaction between the two variables was found which could not be analyzed further because of the nature of the design. In the second place, each S performed on several combinations of tilt and angle, and individual differences between S s were large. No systematic relation of either tilt or angle to precision of linear pursuit movements was, therefore, observed.

* Accepted for publication August 26, 1953. This study was supported in part by the Research Committee of the Graduate School from funds granted by the Wisconsin Alumni Research Foundation.

¹ R. E. Corrigan and W. J. Brogden, The effect of angle upon precision of linear pursuit-movement, this JOURNAL, 61, 1948, 502-510; R. E. Corrigan and W. J. Brogden, The trigonometric relationship of precision and angle of linear pursuit movements, this JOURNAL, 62, 1949, 90-98.

² Angle is designated by Cartesian coördinates: 0° represents the position in which the track is normal to the frontal plane of the body of S , and the arm movement is started with the stylus close to, and continued away from, the frontal plane of the body.

It was hypothesized that tilt is subsidiary to angle from the body in determining precision of linear pursuit movements; that the trigonometric relation of precision and angle of linear pursuit movements will be maintained regardless of tilt, but that its form may be altered as a function of tilt. The present experiment, therefore, was designed to reproduce the trigonometric function of precision in tracking and angle from the body at each angle of tilt selected. In this sense, the trigonometric relation of precision and angle is the dependent variable of the experiment, and angle of tilt is the independent variable.

METHOD AND PROCEDURE

Eight groups of *Ss* each made linear pursuit movements at angles from the body of 0° , 30° , 45° , 60° , 90° , 120° , 135° , and 150° . Every group, however, performed at a different tilt. The tracking surface was tilted for the different groups at angles with intervals of 10° from 0° (horizontal to the floor surface) to 70° . Each increment of 10° in tilt provided a tracking surface which increased the inclination toward the frontal plane of *S* such that at 70° -tilt, the track was most nearly in a vertical plane facing *S*.

Apparatus. The apparatus used in the present experiment is the same as that previously reported.³ It consists of a track formed by two brass plates resting on a piece of glass that *S* traverses with a metal-tipped stylus. Velocity of stylus-movement is controlled by instructions given *S* to match the rate of his movement to that of a small cylindrical target which travels beneath the glass plate of the track at a constant velocity of 3.0 cm. per sec. Control apparatus provides automatically for starting the target, stopping it at the end of the track, and returning it to the starting position where its direction is again reversed for the start of a new trial. The platform on which the track is mounted may be rotated about its center by means of a bearing in the vertical plane. The angle 0° is represented by the track normal to, and target parallel to, the frontal plane of *S*. A disk attached to the bearing of the central vertical axis is notched every 15° to permit rapid and secure selection of the appropriate angle. The platform may also be rotated in the vertical plane by means of a bearing, the surface of which is milled with scale markings every 10° that angle of tilt may readily be adjusted. At 0° -tilt, the track surface is parallel to the floor. Increasing the angle of tilt shifts the plane of the track toward the vertical, with the surface of the track facing the frontal plane of *S*. The stylus and each side of the track constitutes an open switch in parallel with the counting switch of a Potter Electronic Counter, Model 67. Each contact of the stylus with a track-side is registered cumulatively on the counter. When the target is reversed at the end of the track, the input to the counter is closed that no further registration of stylus contact can occur. *E* records the error count for the trial just completed during the time the target is being reversed. When the target reaches the head of the track, the action reversing its direction also clears the counter and registers zero at the start of the next trial.

³ Corrigan and Brogden, *op. cit.*, 61, 1948, 502 f., and 62, 1949, 90 f.

Subjects. The Ss (128 men) were all right-handed and volunteers. None had any previous experience in this experimental situation.

Procedure. The initial step was to seat S in a tank-driver's seat and to fasten a harness about his trunk to keep the right shoulder in a relatively fixed position. The height of the seat and its closeness to the track-platform were so adjusted to body stature that S was both comfortable and able to reach the end of the track with the stylus when the right arm was fully extended. The standardized instructions, reported verbatim in a previous paper, were modified slightly in the present experiment as required by the introduction of tilt as an experimental variable.⁴

The same 8×8 latin square was employed with each of the 8 groups of 16 Ss to control the sequence and ordinal position of the 8 angles. Thus, there were two replications of the randomized square for each of the 8 tilts or groups of Ss. This design eliminates the possibility of an unknown interaction between angle and tilt within the performance of the individual Ss (each S's performance at each of the eight angles was measured at a single tilt of the track surface). S completed 10 practice trials at the initial angle of the sequence to which he had been assigned at random. Thirty seconds after the practice trials were completed, S began the experiment proper. He received 10 trials at each of the 8 angles for his sequence (row of the square) with a 30-sec. rest between each block of 10 trials. Half the Ss were tested by each of the two Es (*GEB* or *RFT*) and with this restriction, S was assigned randomly to *E* and to tilt of the track surface (group).

The same transformation of the error scores for linear pursuit movements used in previous experiments was applied to each score of the present experiment. The natural logarithm was obtained for the sum of the raw score +5 for each S on every trial. Means for the 10 trials in each cell of the row for each S were then computed. These means constituted the measures employed in the analysis.

RESULTS

A summary of the analysis of variance applied to the data appears in Table I, which represents an extension of the usual analysis of a latin square design. Here, not only are angle, ordinal position of angle, and sequence of angle orthogonal variables, but experimenter and tilt of the tracking surface are also orthogonal, both to one another and to the other three variables.

As shown in Table I, tests of significance were made separately for uncorrelated data, between Ss variation, and for correlated data, within Ss variation. Of the total sum of squares, 54.8437, with 1023 degrees of freedom, a between Ss sum of squares of 18.3459 with 127 *df* and a within Ss sum of squares of 36.4942 with 896 *df* were calculated. Further fractionation of each of these sources is shown in the table. The $E \times \text{tilt} \times \text{sequence}$ interaction is the basic error term against which all sources appearing above it in the table were tested. The only *F*-ratios of statistical

⁴ Corrigan and Brogden, *op. cit.*, 61, 1948, 502 f.

significance at the 5-% level of confidence are for tilt of the track-surface and the $E \times SA$ interaction. The significant effect for tilt will be discussed later. That for the $E \times SA$ is not meaningful and no rational explanation for it can be given.

All remaining sources of variations were tested against $RS \times T \times E$ interaction mean square. Here angle, the angle $\times$ tilt interaction, ordinal position, the ordinal position $\times$ tilt interaction, and residual for square are sources of statistically significant variation at the 5-% level of confidence. The significant effect for ordinal position indicates a practice effect from

TABLE I
RESULTS FROM ANALYSIS OF VARIANCE

Source	df	Sum of squares	Mean square	F	F at 5-% level
(1) Total between Ss	127	18.3495			
(a) Experimenter (E)	1	0.0233	0.0233	—	4.08
(b) Tilt (T)	7	4.1266	0.5895	6.03	2.25
(c) Sequence of angle (SA)	7	0.9820	0.1403	1.43	2.25
(d) $E \times T$	7	0.4492	0.0642	—	2.25
(e) $E \times SA$	7	1.6810	0.2401	2.46	2.25
(f) $T \times SA$	49	6.2923	0.1284	1.31	1.69
(g) $E \times T \times SA$	49	4.7951	0.0978	—	—
(2) Total within Ss	896	36.4942			
(a) Angle (A)	7	0.8949	0.1278	5.34	2.05
(b) $A \times E$	7	0.2070	0.0296	1.24	2.05
(c) $A \times T$	49	12.0632	0.2462	10.30	1.42
(d) $A \times E \times T$	49	0.9448	0.0193	—	1.42
(e) Ordinal position (OP)	7	3.3966	0.4852	20.30	2.05
(f) $OP \times E$	7	0.1258	0.0180	—	2.05
(g) $OP \times T$	49	1.6912	0.0345	1.44	1.42
(h) $OP \times E \times T$	49	0.6125	0.0125	—	1.42
(i) Residual for square (RS)	42	1.6108	0.0384	1.61	1.45
(j) $RS \times E$	42	1.2535	0.0298	1.25	1.45
(k) $RS \times T$	294	6.6679	0.0227	—	1.26
(l) $RS \times T \times E$	294	7.0260	0.0239	—	—
(3) Grand total	1023	54.8437			

the first to the eighth of the angles on which the Ss of all groups were tested. This result has been found in earlier studies.⁵ The significant effect for residual for square (square uniqueness) has not been found in any of the previous studies of the trigonometric relation of precision and angle of linear pursuit movements where a latin square design was used. This result suggests that the results obtained with the particular 8×8 latin square selected at random for the present experiment may be unique to this particular square. As will be noted later, however, the trigonometric

⁵ Corrigan and Brogden, *op. cit.*, 62, 1949, 90 f.; G. E. Briggs and Brogden, Bilateral aspects of the trigonometric relationship of precision and angle of linear pursuit-movements, this JOURNAL, 66, 1953, 472-478.

functions fitted to the data of the present experiment are entirely consistent with the results obtained from the earlier studies.

Our primary interest is with the significant effect obtained for angle, tilt, and the angle $\times$ tilt interaction. As was noted earlier, the design of the experiment assigns to the trigonometric relation of precision and angle of linear pursuit movements the function of a dependent variable and to tilt that of the independent variable. Therefore, Duncan's test for differences between ranked treatments was applied separately to the data on the effect of angle for each group of Ss.⁶ The results of these tests indicate that for each set of data, there are at least two groups of scores (*e.g.* most difficult angles and easiest angles) which differed significantly from each other at the 5-% level of confidence. These results justify fitting the trigonometric equation to the data obtained for each of the eight tilts in order to analyze further the significant effect for tilt and the angle $\times$ tilt interaction.

The trigonometric regression equation, $y = a + b \cos 2x + \sin 2x$ was fitted to the empirical data for each angle of tilt by the method of least squares. A plot of each of the eight regression equations is given in Fig. 1. The Chi-square test of goodness-of-fit yielded Chi-squares, for all curves, that are not significant at the 99-% level of probability. The trigonometric equation, therefore, provides a good fit for all tilts. The form of the trigonometric function changes progressively as angle of tilt is increased. These changes are shown more precisely by the curves of Fig. 2. Curve A is a linear plot of the curves of Fig. 1 and shows clearly the progressive changes in the regression equation as angle of tilt is increased from 0° to 70°. Curves B, C, and D present the changes in the parameters of the equation, phase angle, base line, and amplitude, respectively. Since a significant angle $\times$ tilt interaction was found, the effect of tilt upon the trigonometric relation of precision and angle of linear pursuit movements might be expected to be complex. The magnitude of the phase angle does not appear to alter progressively as angle of tilt is increased. Rather, it appears to be of roughly equal size (mean = 17.8°) and positive in sign for angles of tilt of 0°, 10°, 20°, and 30°; of roughly equal size (mean = -20.0°) and negative in sign for angles of tilt of 40°, 50°, 60°, and 70°. Both the baseline and the amplitude, however, show progressive changes in magnitude as the angle of tilt is increased. For each curve, magnitude decreases with increase in angle of tilt to 20°, remains at an approximate minimum from 20° to 50°, and then increases at 60° and 70°.

⁶ R. B. Duncan, A significance test for differences between ranked treatments in an analysis of variance, *Virginia J. Sci.*, 2, 1951, 171-189.

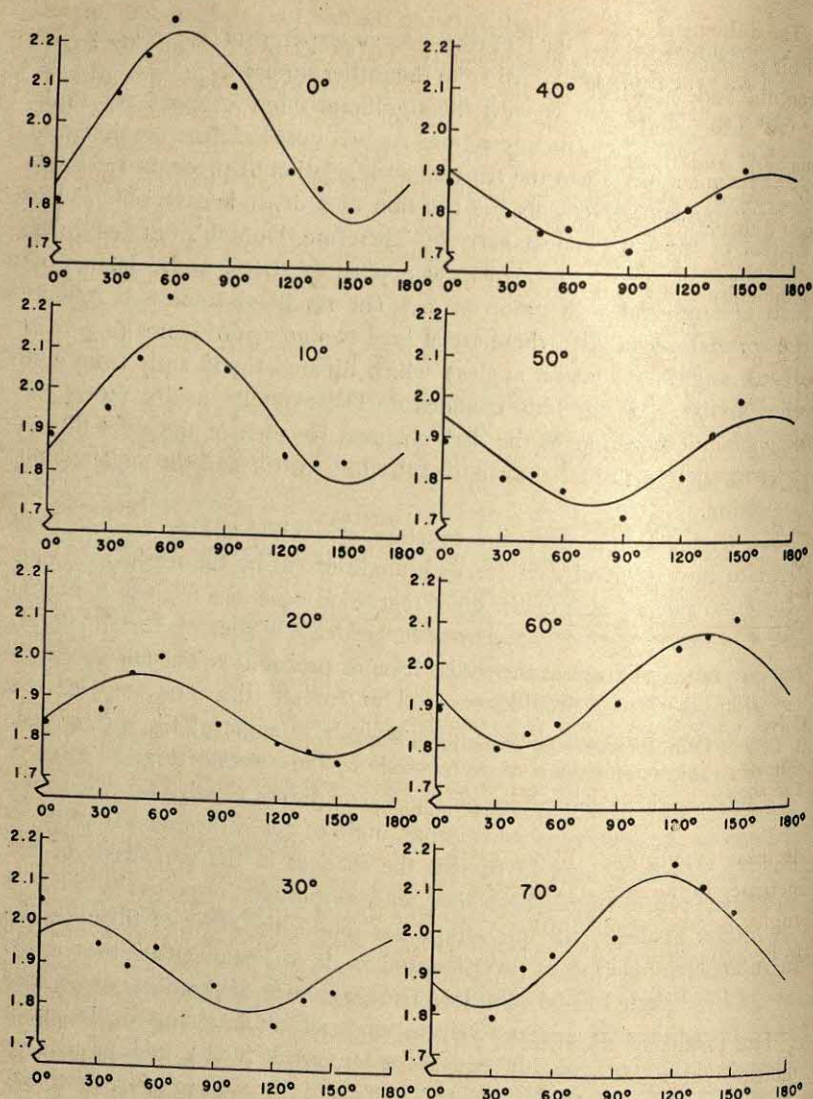


FIG. 1. CURVES OF REGRESSION EQUATIONS FOR PRECISION AND ANGLE OF PURSUIT MOVEMENTS FOR ANGLE OF TILT

For every curve, the ordinate is the logarithm (base e) of the raw error score + 5 and the abscissa represents the angle at which performance was measured. Every point is the group mean of the S s' mean precision of performance for the block of 10 trials at the angle indicated. A separate group of S s was tested at each angle of tilt which is indicated for each of the smooth curves. These curves are plots of the trigonometric regression equation, $y = a + b \cos 2x + c \sin 2x$, fitted to the empirical data by the method of least squares.

The differences in phase-angle between the first four and last four angles of tilt represents virtually a complete reversal in relative difficulty for angle from the body at which the linear pursuit movement is made. For tilts of 0° , 10° , 20° , and 30° , the most difficult angles are those between 30° and 60° , and the easiest between 105° and 135° . At tilts of 40° , 50° , 60° , and 70° , the easiest angles are those between 30° and 60° , and the most difficult are between 105° and 150° .

The changes in baseline and amplitude indicate that the overall level of difficulty and the difference between the most and least difficult angles

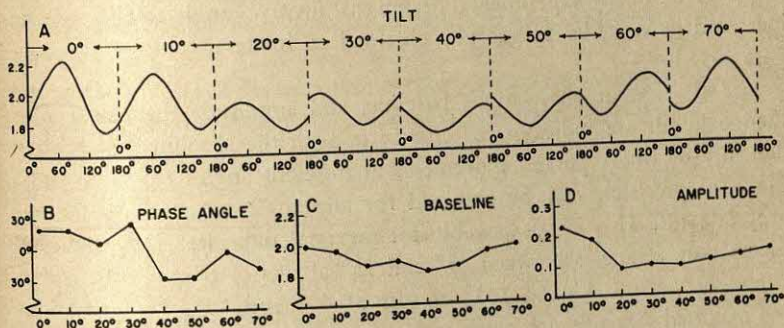


FIG. 2. PHASE ANGLE, BASELINE, AMPLITUDE AND REGRESSION EQUATION AS A FUNCTION OF TILT

Curve A is a linear plot of the regression equations of Fig. 1 against tilt. Curves B, C, and D are plots of the constants of the regression equation against tilt. The ordinate is the transformed error score except for phase angle where it is degrees. The abscissa is angle of tilt, except for A where both angle from the body at which the movement is made and angle of tilt are presented.

decrease as angle of tilt increases to between 20° and 50° and thereafter increase. Difference in difficulty and maximum error for the most difficult angles is greatest at 0° tilt. Next come tilts of 10° and 70° , which appear to be equal; then 60° , with the remaining tilts being about equal.

DISCUSSION

The results of the present experiment indicate clearly that the trigonometric relation between precision and angle of linear pursuit movement is altered as a function of angle of tilt. Unfortunately, the effect of tilt is complex and no single parameter of the trigonometric equation may be used to provide for its quantitative expression. This result is comparable to that obtained by Brogden on the effect of practice on the trigonometric relation of precision and angle of linear pursuit movements.⁷ Phase angle,

⁷ Brogden, The trigonometric relationship of precision and angle of linear pursuit-movements as a function of amount of practice, this JOURNAL, 66, 1953, 45-56.

baseline, and amplitude all were affected by practice, as is the case for angle of tilt. The nature of the effects of practice and tilt, however, are markedly different. The similarity is in the complexity of the effect and the fact that for neither variable is it possible to provide a quantitative expression of the effect as a parameter of the regression equation. These results, however, support the view presented in the introduction of the present paper that angle from the body is the primary determiner of precision of linear pursuit movements and that the integrity of the trigonometric relation of angle and precision of linear pursuit movements should be preserved in any experimental design that involves angle and some other independent variable. The results obtained on bilateral transfer also support this view.⁸

The changes in phase-angle, baseline, and amplitude that occur with change in tilt are indicators of change in both the absolute and relative difficulty of linear pursuit movements at different angles from the body. Maximal precision may be obtained for any angle by choosing the appropriate tilt. This is not possible for maximal difficulty. This condition occurs only at angles between 45° and 75° at tilts of 0° and 10° , and at angles between 105° and 135° at tilts of 60° and 70° . The conditions for maximal difficulty are much more specific than are those for maximal precision. Variables other than tilt, however, do affect maximal and relative difficulty. A shift from right to left hand movements changes the angles at which these maxima occur.⁹ Practice increases the range of angle at which maximal proficiency occurs, but it does not provide for maximal proficiency.¹⁰ The appropriate selection of angle of tilt, right or left hand, and amount of practice, however, provide a fairly wide range of conditions under which maximal precision may be obtained for every angle from the body.

The effect of tilt is undoubtedly a function of concomitant changes in the visual field and in muscular and postural factors. Analysis of tilt in terms of these factors is no more possible than is a similar analysis of angle from the body in such terms. Further investigation into factors influencing precision of linear pursuit movements may best be directed towards the class of variables exemplified by angle, tilt, and the like. The possibility of further analysis in terms of visual and muscular factors, however, should not be ignored.

⁸ Briggs and Brogden, *op. cit.*, 472-478.

⁹ Briggs and Brogden, *op. cit.*, 472-478.

¹⁰ Brogden, *op. cit.*, 45 f.

SUMMARY

The present experiment was designed to provide information about the effect of angle of tilt on a linear pursuit movement. Tilt of the tracking surface was varied by 10° increments from 0° (horizontal) to 70° (near vertical). Linear pursuit movements were made at each tilt at the following angles from the body: 0° , 30° , 45° , 60° , 90° , 120° , 135° , and 150° . An 8×8 latin square design was employed. Two *Es* ran a total of 128 *Ss* at this task. An analysis of variance indicated that angle, tilt, ordinal position, angle $\times$ tilt interaction, ordinal position $\times$ tilt interaction, and residual for square were significant sources of variation. The trigonometric equation, $y = a + b \cos 2x + c \sin 2x$, was fitted by least squares to the data at each angle of tilt. All fits were found to be good by the Chi-square test. The baseline and amplitude coefficients for the trigonometric regression equations decrease in magnitude from 0° to 20° of tilt, are at a minimum between 20° and 50° , and then increase at 60° and 70° of tilt. Two levels of phase-angle were found of roughly equal size but of opposite sign; a mean of 17.8° for angles of tilt of 0° , 10° , 20° , and 30° , and a mean of -20.0° for angles of tilt of 40° , 50° , 60° , and 70° . The effect of tilt cannot be given a quantitative expression in the trigonometric equation. The results of the present experiment and those of earlier experiments on the effect of practice and bilateral transfer lead to the conclusion that the relation of precision and angle of linear pursuit movements is paramount to other experimental variables.

SHIFTS IN PITCH AS A FUNCTION OF PROLONGED STIMULATION WITH PURE TONES

By R. J. CHRISTMAN, Rome Air Development Center, Rome, New York

The Köhler-Wallach study of figural after-effects developed the thesis that the brain functions in a manner somewhat analogous to a volume-electrolyte, and that the experience of any percept, *e.g.* a visual figure, depends upon the dynamic characteristics of the electro-chemical relationships set up by the afferent impulses arriving at the cortical projection-areas.¹ While a number of recent studies have demonstrated phenomena beyond the realm of two-dimensional space perception which might conceivably be explained by the Köhler-Wallach theory,² it may be questioned whether several of these studies should be properly considered as involving figural after-effects, since for a true figural after-effect as described by Köhler and Wallach it is necessary that there be some spatial representation within the central nervous system as a function of some parameter of the stimulus.

Because the perception of auditory pitch does meet the requirement of neuro-spatial localization,³ it should be ideally suited for testing the generality of the Köhler-Wallach theory. If displacements in visual space can be attributed to electrotonic effects within a visual projection area, it is not unreasonable to expect shifts in pitch as a function of electrotonic effects induced in an auditory projection area. According to the writer's interpretation of figural after-effect theory, the pitch of a given tone should shift in a direction away from the pitch of a satiation tone of adequate

* Accepted for publication May 29, 1953. This paper is adapted from a doctoral dissertation submitted to The Ohio State University, June 1952. The writer thanks Professor Samuel Renshaw who sponsored the research.

¹ Wolfgang Köhler and Hans Wallach, Figural after-effects, an investigation of visual processes, *Proc. Amer. Philos. Soc.*, 88, 1944, 269-357.

² R. J. Christman, Figural after-effects utilizing apparent movement as inspection-figure, this JOURNAL, 66, 1953, 66-72; Köhler and Dorothy Dinnerstein, Figural after-effects in kinesthesia, *Miscellanea Psychologica Albert Michotte*, 1947, 196-220; Köhler and D. A. Emery, Figural after-effects in the third dimension of visual space, this JOURNAL, 60, 1947, 159-201; L. Weiskrantz, Figural after-effects in stroboscopic motion, *Quart. J. Exper. Psychol.*, 2, 1950, 113-118.

³ A. R. Tunturi, Audio frequency localization in the acoustic cortex of the dog, *Amer. J. Physiol.*, 141, 1944, 397-403; E. M. Walzl, Representation of the cochlea in the cerebral cortex, *Laryngoscope*, 57, 1947, 778-787; C. N. Woolsey and Walzl, Topical projection of nerve fibers from local regions of the cochlea to the cerebral cortex of the cat, *Johns Hopkins Bull.*, 71, 1942, 315-344.

duration and intensity. Thus, a pure tone which is immediately preceded by a prolonged tone of a slightly lower frequency should appear higher in pitch than the same tone when preceded by silence. Conversely, a pure tone preceded by a prolonged tone of a somewhat higher frequency should appear lower than otherwise expected. A pure tone should not be influenced by a preceding tone of the same frequency.

Apparatus and general procedure. Since the average *O* cannot accurately identify a pitch in absolute terms, it was necessary to use a method of comparison to detect resultant changes in pitch. The method involved subjecting one ear to a satiating tone and then comparing the pitch experienced by stimulating each of the two ears by test-tones. To explain any resulting differences in pitch in terms of figural after-effect theory, it is necessary to assume that the auditory nerves differ in the ratio of contralateral to ipsilateral projection at some neural locus where a spatial representation of the pitch function can occur. On the basis of recent reports by Penfield and Rasmussen⁴ and Rosenzweig,⁵ this assumption appears tenable.

The apparatus consisted of two audio oscillators, matched phones, a decibel meter, and a system of relays and switches which provided automatic as well as manual control of stimulus-duration, phone-selection, oscillator-selection, and loudness at the two ears. The general procedure was to satiate the left ear with a pure tone (satiation-tone) for intervals varying from 1 to 2 min. Then a standard (600~) tone was presented to the left ear and followed immediately by a variable or matching tone (test-tone) to the right ear.⁶ All tones were about 35 db. to minimize contralateral stimulation by bone conduction.⁷ Both a production-method and a modified constant-method were used to determine pitch-matches for the two ears.

EXPERIMENT I

Four experimental conditions were used in Experiment I: (1) standardization (no satiation), (2) 800~ satiation of left ear, (3) 400 or 500~ satiation of left ear,⁸ and (4) control (600~ satiation of left ear).

Procedure. After a 1-min. satiation, or 1 min. of silence in Condition 1, the 600~ standard tone was presented to the left ear for 2 sec. After a 2-sec. pause, a test-tone of 400, 500, 700 or 750~ was presented to the right ear. The order of these basal settings was varied in a systematic manner which was not apparent to the *O*, who proceeded to adjust from the basal setting to a point of subjective equality with the preceding standard. Four matches were made after each satiation. Each of the 5 *O*s, graduate students in psychology, made 24 matches under each of the 4 experimental conditions—a total of 96 matches for each *O*. All *O*s were

⁴ Wilder Penfield and A. T. Rasmussen, *The Cerebral Cortex of Man*, 1950, 153.

⁵ M. R. Rosenzweig, Representation of the two ears at the auditory cortex, *Amer. Psychologist*, 5, 1950, 270.

⁶ A sufficient number of *O*s were used to minimize any possible effects due to diplacusis binauralis.

⁷ S. S. Stevens and Harlow Davis, *Hearing: Its Psychology and Physiology*, 1938, 215.

⁸ Of the 5 *O*s, 3 had 400~ satiations and 2 had 500~ satiations.

limited to one satiation per session and to two sessions per day to avoid any cumulative effect.

Results. An analysis of variance was performed on the data with results shown in Table I. Variances due to conditions, to *O*s, and to basal settings are significant beyond the 1-% level. The interaction between conditions and *O*s was statistically significant (beyond the 1-% level), a result which indicates individual differences in the degree of susceptibility to satiation effects as a function of the frequency of the satiating tone. This interpretation is supported by the data for individual *O*s presented in Table II. Also

TABLE I
VARIANCE ANALYSIS FOR EXPERIMENT I

Source	Sums of Squares	df	Est. var.
conditions	23,799	2	11,899
<i>O</i> s	14,088	4	3,522
basal settings	66,477	3	22,159
conditions $\times$ <i>O</i> s	30,423	8	3,803
<i>O</i> s $\times$ basal settings	127,440	12	10,620
conditions $\times$ basal settings	1,049	6	175
conditions $\times$ <i>O</i> s $\times$ basal settings	18,061	24	753
remainder	101,204	300	337
Total	382,541	359	

of interest is the magnitude of the variance due to the interaction of *O*s and basal settings.

The significance of the variance due to conditions indicated the advisability of a further examination of the data and the application of separate *t*-tests. While 24 of the 25 intra-*O* comparisons give differences in the predicted direction, several *O*s show significant shifts as a result of only one of the two satiation conditions expected to result in pitch-displacements. For the mean settings there is no significant difference between the standard condition and the 600~ control condition, as would be predicted. The *PSE*s of 595.8 and 595.3 are for all purposes identical ($t = 0.18$). All five mean differences involving 800 and 400 or 500~ satiations are in the direction predicted by the original hypothesis, and all are significant beyond the 1-% level of confidence, with *t*s ranging from 3.01 to 7.47.

Since four matches were made after each satiation, it was possible to examine the effect of time between satiation and testing on the magnitude of the displacement. For the 800~ satiation, the mean *PSE*s were 578.2, 585.6, 586.0, and 586.2~ over a time-interval of about 30 sec. For the 400 and 500~ satiation, the mean *PSE*s were 608.2, 599.3, 606.5, and 601.1~. In both cases the general trend is in the direction of a decreasing

effect with time, results which are in accord with those of Hammer for visual figural after-effects.⁹

EXPERIMENT II

A modified constant-method was used in Experiment II, involving a 600~ standard presented to the left ear which was followed by a variable tone to the right ear. The frequency of the variable, which was judged to be *high* or *low* relative to the standard, was 590, 595, 605, or 610~.

Procedure. Two-category judgments were used throughout, and whenever *O* could not make a judgment—a rare event—the same comparison-pair was repeated at a later point in the series. Except for the standard condition, in which 20 judgments constituted a session and no satiation was given, each experimental session consisted of 10 judgments with each comparison-pair preceded by satiation of the left ear. The duration of the standard tone, of the interval, and of the variable was 2 sec.,

TABLE II
PSEs FOR EACH *O* UNDER EACH CONDITION OF EXPERIMENT I

<i>O</i>	Standardization	800~	400 & 500~	600~
A	595.2	590.8	595.3	596.3
B	586.0	549.8	614.0	591.2
C	597.0	590.0	597.3	593.3
D	602.0	592.9	607.7	600.0
E	598.7	596.3	604.7	595.5
Mean	595.8	584.0	603.8	595.3

the three intervals being automatically regulated. The four variable stimuli were paired with the standard in a randomized order.

In addition, to three control conditions designed to measure constant errors due to the equipment, a total of 21 experimental conditions were used. These conditions, involving 10 satiation frequencies, each for 2 durations, are listed in Table III. Five *O*s were used under each experimental condition, yielding a total of 50 judgments per condition and 100 for the standard condition. In the experiment as a whole, 120 *O*s made 1300 judgments, all but 200 being preceded by 1- or 2-min. satiating periods.

Results. Due to the small number of judgments and the restricted range of the variable stimuli necessitated by the large number of experimental conditions, it was impossible to examine the data by conventional methods. As an alternative procedure, an analysis was made of the number of *high* and *low* judgments under each of the 21 conditions. Of the 40 *O*s satiated with pitches below 600~, only 9 had a majority of *high* judgments, while 25 judged the variable tone to be *low* more often than *high*. This departure

⁹E. R. Hammer, Temporal factors in figural after-effects, this JOURNAL, 62, 1949, 337-354.

from chance expectancy is significant beyond the 5-% level of confidence. For the 50 *Os* satiated with tones above 600~, the results are even more significant (beyond the 1-% level). Only 6 of these *Os* gave a majority of *low* judgments while 41 gave results indicating pitch-shifts in the expected direction.

Table III shows the percentage of *high* judgments under each of the 21 experimental conditions. A preponderance of *high* judgments indicates a depressing effect induced by the satiating tone, while a value smaller than 50% *high* judgments indicates that the pitch at the left ear has been elevated. Satiation-tones below 600~ apparently produced an elevation of pitch, while satiation-tones above 600~ produced depression. For the standard condition, the 100 judgments are divided 52 to 48. For the 600~ control condition there is no evidence of any displacement effect at 1 min., but the percentage of judgments *high* at 2 min. seems too great to be due to chance variation, a finding which will be considered later.

The influence of particular satiating frequencies cannot be determined from the data in Table III due to the small number of *Os* tested under each

TABLE III
PERCENTAGE OF HIGH JUDGMENTS FOR EACH CONDITION OF
EXPERIMENT II

Satiation frequency	Duration (sec.)	
	60	120
500	50	16
550	42	50
580	40	32
590	40	36
600 (Control)	52	70
610	64	31
620	64	84
650	80	78
700	66	58
800	74	88
Standardization		52
500 to 590 incl.	43	34
610 to 800 incl.	70	74

condition (5 *Os*, 50 judgments). If optimal frequencies exist, this experiment did not discover them, although it is possible, of course, that they may lie beyond the range studied. The data of Table III suggest, however, that the higher frequencies have a greater effect than the lower. It may also be observed from Table III that a greater effect is produced by a 2-min. period of satiation than by a 1-min. period, a finding which holds both for high and for low satiating tones.

PSEs were calculated by a graphic method for several combinations of conditions. The *PSEs* for the low satiation tones, the standard condition, the control condition, and the high satiation-tones were 603.0, 597.2, 597.4, and 590.0 respectively. These values compare very favorably with those found directly in Experiment I (Table II). The relatively large discrepancy between the two experiments with regard to satiating tones above 600~ (590.9 as compared with 584.0) is explainable on the basis of the limited range of variables used in Experiment II. Since the lowest variable tone was 590~, the *PSE* was largely determined by this value.

DISCUSSION

The two experiments here reported provide strong evidence of shifts in pitch produced by prolonged stimulation with pure tones. When a given pitch is preceded by a satiating tone of a somewhat lower frequency, its pitch is slightly elevated and the opposite effect is produced by a satiating tone of higher frequency. The fact that a greater shift seems to be produced by the higher satiating tones might be explained in several ways:

(1) A pure tone is not transmitted through the entire course of the pre-labyrinthine mechanism of the ear without distortion. On the contrary, numerous intra-aural harmonics appear to be generated, probably in the middle ear, and these harmonics may very well serve as satiating tones in exactly the same way as does the fundamental. Since the intra-aurally induced tones would be predominantly higher in frequency than the fundamentals from which they were generated, the effect produced by a low tone would be a function of both the low fundamental of relatively high amplitude and the higher-frequency harmonics which would tend to act in opposition to the fundamental. For satiations with tones higher than the standard, the harmonics would tend to reinforce the fundamental, thus increasing the effect of the satiating tone.

(2) The possibility of a constant difference due to the apparatus which presented the stimuli to the two ears must not be overlooked. Results obtained under two standardization conditions, however, make this explanation untenable. A reverse standardizational series gave some evidence of a constant difference opposite to that actually found under satiating condition, which indicates that the greater shifts produced by high tones cannot be attributed to equipment.

(3) The difference in question might be explained as a negative time-error running through the experiments. This explanation is faced with two difficulties: (a) there is very little evidence in the literature for a reliable time-error in pitch—what evidence exists is fragmentary and highly contro-

versial; and (b) time-errors are themselves empirical phenomena for which no entirely adequate theory is presently available. Indeed, it seems that the displacement-effects found in this study might provide the basis for a new explanation of time-errors, rather than the converse.¹⁰

The apparent shift produced by a 2-min. period of satiation with the 600~ tone may provide further support for the explanation. In neither experiment was there any evidence of a shift in pitch after a 1-min. satiation with the standard. The difference may be accounted for on the assumption that the harmonics were sufficiently weak that their effects would be measurable only after a longer period of satiation.

Two instances of agreement with data on visual after-effects seem pertinent to any discussion of the basis for shifts in pitch here reported. As in Hammer's results for vision, the magnitude of the after-effect appeared to increase with the duration of the satiating stimulus.¹¹ Such a result, it should be noted, would not be predicted on the basis of an explanation involving 'unconscious comparison' or 'contrast effect.' There is also some evidence, as in Hammer's study, for a diminution of the after-effect with time. While these results do not prove that the auditory effect is neurologically comparable to visual effects, there is nothing to contradict this assumption.

While the results here reported seem to be best accounted for in terms of the Köhler-Wallach theory, the possibility that they may be traced to cochlear functions cannot entirely be ignored. One reason for assigning the effect to a higher level than the cochlea, however, is the nearly complete absence of any well substantiated evidence for cochlear fatigue or adaptation in pitch perception. In all the literature reviewed, the only mention of shifts in pitch as a function of prolonged antecedent pure-tone stimulation was found in a report by Békésy.¹² Neither that study nor a subsequent one by Egan and Meyer,¹³ in which pitch shifts were produced by simultaneous stimulation with a 90-cycle band of noise, are conclusive in locating the effect at the cochlear level.

It might be possible to explain the data of the present study in terms of a

¹⁰ The as yet undeveloped hypothesis would involve three factors: (1) fading trace, particularly in intensive dimensions; (2) satiation-produced displacements as described in this report; (3) subjective factors such as set and attitude.

¹¹ Hammer, *op. cit.*, 348-351.

¹² George von Békésy, Zur Theorie des Hörens: Über die Bestimmung des einen reinen Tonempfinden entsprechenden Erregungsgebietes der Basilarmembran vermittels Ermüdungserscheinungen, *Physik. Zsch.*, 30, 1929, 115-125.

¹³ J. P. Egan and D. R. Meyer, Changes in pitch of low tones as a function of the pattern of excitation produced by a band of noise, *J. Acoust. Soc. Amer.* 22, 1950, 827-33.

more conventional picture of neural functioning, as Osgood and Heyer did in the case of visual after-effects.¹⁴ Not only, however, does that explanation offer little improvement over the volume-electrolyte theory, as Smith points out,¹⁵ but there is, in addition, some reason to doubt its applicability to the auditory data, since the presence of slight eye-movements necessary for the statistical theory have no known counterpart in the perception of pitch. The cochlea is not known to possess any minute scanning properties, although it is possible that slight periodic fluctuations in the cochlear or cortical locus of the pitch-function may occur. It seems that the effect reported here can be accounted for in terms of an electrotonic effect in a volume electrolyte, an effect such that the neural locus of a given pitch will be shifted away from an area previously affected by a prolonged satiation tone. In contradistinction to the limitations set forth by Köhler and Wallach,¹⁶ it would appear that the shifts in pitch might very well be isomorphic with actual areal displacements rather than a function of increased resistance or impedance between the neural loci of the standard and the variable stimulus.

While the presented findings are based on an extremely restricted range of auditory frequencies, there is no particular reason to expect different results for any other frequencies which might be shown to have some degree of neuro-spatial localization. It is quite possible, of course, that for higher frequencies (above 1600~) no displacement effects would be found.¹⁷ Further work utilizing a greater range of frequencies might well be undertaken.

SUMMARY

In two experiments shifts in the pitch of a standard tone of 600~ were found following monaural stimulation (satiation) with pure tones of 450-800~ and of 1- or 2-min. duration. The pitch of the standard was lowered by satiating tones of higher frequency and raised by satiating tones of lower frequency, the effect being somewhat greater in the former case than in the latter. The magnitude of the effect varied directly with the duration of the satiating tones and inversely with the time between satiation and testing. The results of the experiments can be explained in terms of the Köhler-Wallach theory.

¹⁴ C. E. Osgood and A. W. Heyer, Jr., A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1952, 98-118.

¹⁵ K. R. Smith, The statistical theory of the figural after-effect, *ibid.*, 59, 1952, 401-402.

¹⁶ Köhler and Wallach, *op. cit.*, 337.

¹⁷ Tunturi (*op. cit.*, 400-402), found cortical loci on the middle ectosylvian gyrus of the dog for frequencies only up to 1600~.

BINOCULAR SUBJECTIVE COLORS AND THE VISUAL MECHANISM

By ROBERT W. BURNHAM, Rochester, New York

A description will be given here of unusual, perhaps unique, chromatic visual experiences, which may provide a clue to at least one aspect of the functioning of the visual mechanism. These events may be categorized as subjective colors because they seem more indicative of the visual mechanism than of the visual stimulus. The term 'subjective color' is redundant in one sense, but traditional usage may justify an old term in a situation where new terms might complicate the problem of communication.

These subjective colors are introspectively similar to what have been described as Fechner's colors,¹ but require a basic difference in the stimulating conditions to bring them about. Fechner's colors may be seen, as a review of the literature shows, by monocular as well as binocular vision.² Other subjective colors, apparently related to Fechner's colors, may be seen either monocularly or binocularly when a diffuse light source is pulsed at comparably slow rates. The subjective colors presently to be described may be seen in a similarly pulsed system, but only at much higher pulse-rates; at high luminances of the pulsed source; and in binocular, *but not monocular*, vision. It is principally the fact that these subjective colors can be seen only in binocular vision that sets them apart from Fechner's colors, and provides a clue to the functioning of the visual mechanism.

There has been some question as to what parts or activities of the visual mechanism are crucially responsible for subjective colors since such color appearances do not seem to bear any direct relationship to the spectral distribution of the incident radiation. Gehrke, for example, illuminated various black-and-white disks with a narrow spectral band sodium source, and when a variety of hue-responses was reported he concluded that activity of a 'central' mechanism was required to explain his results.³ Hess on the other hand, attempted to provide evidence for a retinal interpretation of Fechner's colors by breaking up the usual monocular stimulus-pattern into two parts which could be fused haploscopically.⁴ When no subjective colors were seen, it was concluded that a 'central' interpretation was

* Accepted for publication July 22, 1953. From the Color Technology Division, Eastman Kodak Company.

¹ S. M. Newhall, *Psychological Concepts: Sensory Aspects of Color*, Ch. IV in *The Science of Color*, Committee on Colorimetry of the Optical Society of America, 1953, 109 ff.

² Jozef Cohen and D. A. Gordon, The Prevost-Fechner-Benham subjective colors, *Psychol. Bull.*, 46, 1949, 97-136.

³ E. Gehrke, Neue Versuche über Farbensehen, *Ann. Physik*, (6) 2, 1948, 345-354.

⁴ E. H. Hess, 'Subjective' colors: Retinal vs. central origin, this JOURNAL, 65, 1952, 278-280.

doubtful, although not entirely precluded. Even when the entire pattern was viewed by one eye and the colors were seen, it can only be said that a whole chain of physiological events preceded the perception. Neither Gehrke nor Hess, with this purely stimulus-response kind of evidence, isolated any particular anatomical level which could be said to be crucially responsible for the perceptions.

Isolation of levels is difficult, if not impossible, when a monocular stimulus initiates a chain of events ending in a visual response. When, however, stimulation of both eyes is required for a particular perception, which cannot be experienced upon stimulation of either eye alone, it may be deduced that the perception is crucially mediated by a physiological interaction which can only take place at or beyond the first anatomical position where interaction is possible. Although such observations may not pin-point a particular interaction, they at least narrow the search by eliminating levels below which the interaction could *not* take place. Also, with some knowledge of the kinds of physiological activity to be expected at or beyond the crucial level, important deductions may be made concerning the kinds of activity leading to the perception.

The observations to be reported here point to a level in the hierarchy of the color mechanism below which the visual effects cannot be accounted for. The present study, however, by no means constitutes the *experimentum crucis* for a central *vs.* retinal hypothesis of all subjective colors, for it remains to be established how closely related other subjective colors are to those described here.

This report is concerned with a description of two subjective colors which were observed under restricted, yet specifiable conditions, and could be seen *only* when both eyes of an observer were appropriately stimulated. These colors may be described as a striking red and blue which, relative to Fechner's colors, are of unusually high saturation and brightness.

The description to be given may not constitute a new discovery. There are several reports in the literature which conceivably concern the same phenomenon, although it is uncertain whether both eyes were required for the perceptions reported. Also, the conditions under which the present colors were reported are different enough to suggest that the older reports may concern other (though possibly related) phenomena.

Edridge-Green reported in 1895 that he had verified certain observations described earlier by Helmholtz in his 1866 Edition of the *Handbuch der Physiologischen Optik*. Edridge-Green said that "if a disc, marked with black-and-white sections, be rotated with a certain rapidity, the field appears to be covered with a pattern composed of hexagonal spots. . . . Each of the hexagonal spots is . . . surrounded by a red thread, which appears to be moving in minute drops. This field seems to be pervaded by a greenish hue, which flows . . . (toward the center of the field)."⁵ In a later study, Pauli and Wenzl reported that, at speeds of 33 to 40 c.p.s., red and green spots of moderate saturation appeared *irregularly* near the rim of a

⁵ F. W. Edridge-Green, The artificial spectrum top, *Nature*, 51, 1895, 321. *

rotating sector disk.⁶ These authors report a red and green, instead of the red and blue found in the present study, but their descriptions sound enough like the present observations that they cannot be ignored as possible earlier reports of the same thing.

PROCEDURE AND OBSERVATIONS

The colors were first seen by the author in the course of a systematic introspective survey of the chromatic and articulatory phenomena to be observed in flickering fields. A number of parameters were investigated over a wide range. There were 286 combinations of conditions used, which included variation in the luminance of the pulsed source (7.2 to 1540 footlamberts), chromaticity of the source (color temperatures from 2580° to 3700°K, and 3700° filtered by Kodak Color Compensating Filters CC50 red, yellow, green, cyan, blue, and magenta), pulse-rate (8 to 43 c.p.s.), observing time (up to 15 min. per condition), and field-size (2° and 40°). Comparisons were also made between monocular and binocular viewing under each combination of conditions. Different pulse-rates were produced by varying the speed of rotation of a sector disk placed in the beam of a Master Model Kodaslide Projector (1000-wa.), and the light-dark intervals were maintained at a one-to-one ratio. Various luminance levels were attained by attenuating the projector beam with an iris diaphragm. Field-size was controlled by projecting the image of fixed apertures located in the slide holder of the projector. The pulsed beam was projected to the center of a 24-in. square translucent opalized glass screen which was viewed by *O* from the opposite side. Ambient illumination on the glass screen beyond the test field, and in the rest of the room, was at about the threshold between photopic and scotopic levels.

With appropriate variation of the flicker conditions it was possible to reproduce many of the subjective forms and figures reported by others, such as the "Lichtschattenfiguren" of Purkinje;⁷ the achromaticness of a flickering, chromatic, peripheral field reported by LeGrand and Geblewicz;⁸ the floating "ubiquitous violet cloud" reported by Gebhard,⁹ Carrière,¹⁰ and others;¹¹ the relatively stable violet flecks on a yellow-green ground reported by Brown and Gebhard for the unstimulated eye in monocular viewing;¹² the chromatic appearances induced in black-and-white pictures as reported by Stewart;¹³ and a succession of colors attributable to the illuminant and its after-image reported by many others.¹⁴ As the observations

⁶ R. Pauli and A. Wenzl, Über Farbenempfindungen bei intermittierendem farblosem Lichte, *Arch. f. d. ges. Psychol.*, 48, 1924, 470-484.

⁷ J. E. Purkinje, Beiträge zur Kenntniss des Sehens in subjectiver Hinsicht (1819), reprinted in his *Opera Omnia*, 1, 1918, 6-9.

⁸ Y. LeGrand and E. Geblewicz, Recherches sur la vision latérale, *Rev. Opt. (théor. instrum.)*, 17, 1938, 257-274.

⁹ J. W. Gebhard, Chromatic phenomena produced by intermittent stimulation of the retina, *J. Exper. Psychol.*, 33, 1943, 387-406.

¹⁰ Y. Carrière, Virage du blanc au violet par vision stroboscopique, *J. Phys. Radium*, 9, 1948, 122-129.

¹¹ See the review by Cohen and Gordon, *op. cit.*, 97-136.

¹² C. R. Brown and Gebhard, Visual field articulation in the absence of spatial stimulus gradients, *J. Exper. Psychol.*, 38, 1948, 188-200.

¹³ G. N. Stewart, Color phenomena caused by intermittent stimulation with white light, *Amer. J. Physiol.*, 69, 1924, 337-353.

¹⁴ Cohen and Gordon, *op. cit.*, 97-136.

continued, however, unique and striking subjective colors appeared under relatively restricted conditions which could not be found under other conditions.

Prior to the appearance of the red and blue colors with which this report is principally concerned, the author, serving as *O*, had consistently reported the fluctuating, moderately bright, moderately saturated violet cloud surrounding a bright, unsaturated, greenish-yellow central (foveal) area which had been reported by others.¹⁵ The ubiquitous violet cloud was seen over a considerable range of color temperatures and pulse-rates, for short and long viewing intervals, for both small and large fields, and for both monocular and binocular viewing. When, however, the conditions were narrowed to give pulsed luminances of 600 to 650 footlamberts, at pulse-rates of 37 to 43 c.p.s., for the 40° field, for viewing intervals from roughly 10 sec. to 2 min., and for binocular viewing (only), new and striking subjective colors appeared which could not be found beyond the limits set by these conditions. Their appearances was, however, independent of color temperature, and remained unaltered, with each of the six moderately chromatic filters in the beam. *O* reported that superimposed on the violet cloud was a large area of a brilliant red, judged to be equivalent in appearance to Munsell 5R 6/12 (viewed under Macbeth daylight), and a comparably large area of a brilliant greenish blue, judged as 10B 6/8. These two colors seemed to occupy the same or overlapping large central areas at what could be the same or successive moments. They seemed to fluctuate in depth so that one color could be described as being seen through the other, although either color could be seen in either position. The colors appeared in a shimmering form of the volume mode of appearance, and had ever-changing indefinite edges. Each was overlaid almost always with a checkered pattern of white pinpoints or stars. The colors were so striking and obvious, and so unexpected, that further study seemed desirable.

The appearance of both colors was verified by several uninstructed *O*s, and observed again and again by the author. It was found that casual viewing without rigid fixation, accompanied by frequent forced (or involuntary) blinking and eye movements, with occasional 'rest' acquired by looking at the darker surround, made it possible to continue to see the colors for an extended period of time. Persistent fixation, without eye movements, on the bright flickering field resulted in the disappearance of the colors after several minutes. The colors usually appeared first after about 5 to 10 sec. of viewing. One *O* was unable to see the colors in a number of attempts. Two others did not see the colors at first but in subsequent sessions reported them immediately and unambiguously. The colors are so vivid and their areas so large that they cannot easily be overlooked when they do appear.

The fact that they were seen only in binocular vision is particularly interesting, since other subjective colors may be seen with either eye alone as well as binocularly. This binocular requirement was verified many times when, by closing one eye or the other while the colors were present, the colors always disappeared and left a neutral gray flickering pattern, or the violet cloud, or both. Also, by fixating beyond the test-field, that double vision of the circular flickering field was attained, the colors persisted up to the edge of the binocular portion of the field but cut off sharply at the boundaries of the binocular and monocular parts of the field, and never appeared in the monocular parts.

¹⁵ Cohen and Gordon, *op. cit.*, 97-136.

It was possible that closing one eye decreased the effective luminance of the field, and that monocular luminance comparable to that previously found in the binocular field might be required for appearance of the colors. Consequently two types of check observations were made. In the first, as one eye was closed the monocular luminance was simultaneously increased to a level approximately equal to that required binocularly for appearance of the colors. In the second, the luminance of the test-field was present at the high level, one (or the other) eye was covered with an opaque patch, and monocular observations alone were made. In neither situation, with repeated series of observations, did the colors appear. These observations supported the notion that simultaneous external stimulation of both eyes is required.

As a means of getting some idea of the generality of the phenomenon, a flicker-field was devised which utilized a General Radio Strobolux 648-A as the variable pulsing source. The Strobolux was triggered by a General Radio Strobotac 631-B. The argon lamp of the source was of considerably higher color temperature (about 9000°K) than any used in the first series of observations. Its luminance could not be estimated accurately but it was at a high level. The light-dark intervals had approximately a one-to-four ratio when the Strobotac was set at 40 c.p.s., as compared to one-to-one in the earlier observations. A gray cardboard screen of about Munsell N5/ was used as a surround in which a circular hole was cut that subtended a visual angle of 18° at the *O*'s eye. Variable illumination of the surround from 0.18 to 13 footlamberts was obtained with a bank of four Macbeth daylight (9000°K) lamps. Pulse-rate was varied, in steps of about 2 c.p.s., from 30.1 to 60.0 c.p.s.

Four *O*s were used who in preliminary observations had been able to see the two colors quickly and with no question. One *O*, the author, made six series of observations and each of the other three *O*s made one series of observations. A series consisted of one observation at each of 11 pulse-rates at each of seven levels of surround illumination. *O* was seated before the field and given 5 min. to adapt to the 0.13 footlambert surround level. The flickering field was then exposed for 15 sec., after which it was covered with a piece of the opaque surround material for 30 sec. During the 30-sec. interval after the exposure, the observer reported 'yes' or 'no' concerning the appearance of the distinctive red and blue colors. He was instructed to ignore the violet cloud which usually was also apparent. In addition, he estimated, where possible, the proportion of the 15 sec. during which the red and the blue colors were present. Responses were recorded by *E*. At the end of the 30-sec. interval, during which the surround was viewed, the field was again exposed for 15 sec. at a different pulse-rate, but at the same surround luminance level, until all 11 pulse-rates had been used at that level. Then the surround luminance was changed, *O* readapted for 3 min., and the observations were continued. Pulse-rate was randomized for the exposures at each surround luminance level, but luminance levels of the surround were varied in order from low to high or from high to low to permit progressive adaptation to the new level. In the case of the author who made six series of observations, half were made with luminance levels progressing from low to high and half in the reverse order.

The frequencies of the 'yes' responses, indicating the appearance of the colors, are shown in Table I. It may be seen that they center around 40 c.p.s., and that

the range of effective pulse-rates decreases as surround luminance level is increased. That temporal part of each pulsed interval during which the blue and red were present varied greatly from one observation to another and among *O*s. There was no consistent predominance of one color over the other.

In addition to the systematic observations by these four *O*s, all of whom could see the colors within a few seconds after the field was exposed, seven other *O*s were asked to look at the 18° field, with the surround at the 0.13 footlambert level and with the pulsed source at 40 c.p.s., and report what they saw. This combination of conditions was employed because the more systematic observations reported

TABLE I
FREQUENCY WITH WHICH THE RED AND BLUE COLORS WERE SEEN

Each entry in the table could have a maximal frequency of 9 if the colors had been seen at all pulse-rates and surround luminance-levels

Pulse-rate of test-field (in c/s)	Surround luminance in footlamberts							Total
	0.13	1.4	3.8	8.3	10.5	11.5	13.0	
60.0	0	0	0	0	0	0	0	0
56.6	0	0	0	0	0	0	0	0
53.2	1	0	0	0	0	0	0	1
50.0	1	1	1	1	0	0	0	4
46.7	4	2	1	1	1	1	0	10
43.4	7	3	2	2	1	2	0	17
40.0	9	9	6	7	2	1	0	34
36.7	3	3	2	1	0	0	0	9
33.3	2	0	0	0	0	0	0	2
31.7	0	0	0	0	0	0	0	0
30.1	0	0	0	0	0	0	0	0
Total	27	18	12	12	4	4	0	77

above indicated a greater likelihood that the colors would be seen most frequently under these conditions. The *O*s varied in the amount of information they had concerning what they were expected to see. Five *O*s reported the colors almost immediately and unambiguously. Two *O*s, however, who observed a number of times over a period of several months, have been consistently negative in their reports. One of these *O*s was the one who could not see the colors under the earlier projector conditions. There is little likelihood that the colors were present but ignored. When the colors appear, they are so very prominent in the field, and so striking as colors, it seems impossible that they would be overlooked. All the *O*s had normal color vision as determined by AO plates and the Nagel anomaloscope.

The necessity for using both eyes in order to see the colors was verified by all of the *O*s who could see the colors at all. This was done both by closing one eye at a time, and by fixating beyond the field to give a part binocular, part monocular, field as in the earlier observations.

The author again was able to see the same blue and red colors when viewing the flickering field through Kodak CC50 red, yellow, green, cyan, blue and magenta filters. In spite of a hue-change in the flickering illuminant the same red and blue were seen.

The author also observed, on moving his head toward the field as close as he

could (until his visual angle was about 150°), that the colors appeared much less frequently and remained in the central portion of the field in an area subtending an estimated angle of 20° . In this case, of course, the 'surround' flickered and was of the same brightness as the test-field since the surround and test-field actually constituted a single continuous field.

DISCUSSION

Brilliant red and blue perceptions are elicited by a binocular field of possibly any spectral composition, provided the luminance is relatively high and the field is pulsed at about 40 c.p.s.

These binocular subjective colors have implications for the functioning of the visual mechanism. The mechanism must function so as to account for a binocular response which is qualitatively different from that observed when either eye alone is pulsed by the same stimulus, and which is independent of the spectral composition of the stimulus. Such a mechanism is found in the Troland-Fry modulation hypothesis where variation in the spectral composition of the stimulus produces differential modulations of the optic nerve current.¹⁶ The differential modulations are, in turn, associated with varying impressions of chromaticness. In particular, Fry's explanation of binocular color mixture applies here.¹⁷ When the two eyes are differently stimulated, as in binocular color mixture, Fry assumes that unlike modulations from the two retinas interact in such a way as to produce a third kind of pattern in the common cerebral paths.

To account for unlike modulation patterns in the present case, it need only be assumed that the two eyes are responding with some slight difference to the same stimulus. This does not seem unreasonable where a rapidly pulsed stimulus is used. The response difference could be one of frequency, phase, or whatever, but such a difference, plus resolution of the difference into a third pattern in the common cerebral paths, could account (as in binocular color mixture) for a perception which is different with binocular stimulation than with stimulation of either eye alone. The resolved nerve pattern could conceivably be a 'beat' pattern, in analogy with the auditory situation where the simultaneous action of two slightly different sound frequencies produces a third pattern. The color perceptions would, then, be more related to the 'beat' pattern than to the spectral quality of the stimulus. These subjective colors certainly bear some relationship to the

¹⁶ L. T. Troland, The enigma of color vision, *Amer. J. Physiol. Optics*, 1, 1920, 317-337; 2, 1921, 23-48; and G. A. Fry, A photo-receptor mechanism for the modulation theory of color vision, *J. Opt. Soc. Amer.* 35, 1945, 114-135.

¹⁷ G. A. Fry, Binocular integration of hue and brilliance, *Arch. Ophthalm.* 15, 1936, 443-456.

pulse-rate of the stimulus, since they are only seen when the stimulus is pulsed at about 40 c.p.s.

This speculation concerning differential response can, fortunately, be subjected to experimental test. Such a test would require the use of Sherrington's flicker technique so that each eye can be independently stimulated, and as monocular variations in frequency and phase are provided.¹⁸ Assuming frequency of stimulation to be the appropriate variable, it may be predicted that *Os* who see the colors when both eyes are stimulated at the same (40 c.p.s.) frequency, should not see them if the frequency of the pulsed source is varied sufficiently in one eye as compared to the other. In other words, it might be possible to introduce artificially the same modulated response in two eyes that normally respond with slight differences to identical pulse stimuli. Conversely, *Os* who do not see the colors when both eyes are stimulated alike, should see them if the frequency of a pulsed stimulus is appropriately varied for one eye as compared to the other.

It may be, of course, that frequency is not the appropriate variable. Perhaps the monocular responses are slightly out of phase, in which case appropriate monocular phase differences in the pulsed stimulus should be effective in eliminating the colors in the case of persons who normally see them, or in producing the colors for people who do not normally see them. These are at least clues to experimental approaches which could be made to a better understanding of the mechanism involved.¹⁹ It may also be true, as Bott has suggested, that other colors can be found by variations of this binocular procedure.²⁰

SUMMARY

Introspective observations are reported on two possibly unique subjective colors which may be described as a brilliant red and a brilliant greenish-blue. The colors seem unique principally because they can be seen only in binocular vision, never, apparently, in monocular vision. Implications for functioning of the visual mechanism are discussed, and a testable 'beat' hypothesis is developed.

¹⁸ C. S. Sherrington, On binocular flicker and the correlation of activity of "corresponding" retinal points, *Brit. J. Psychol.*, 1, 104-105, 26-60.

¹⁹ A verification of the 'beat' hypothesis would provide evidence against theories of binocular vision which question the possibility of central fusion or interaction. See G. L. Walls, Is vision ever binocular? *Opt. J. Rev. Optom.*, 85, 1948, 33-43, and F. H. Verhoeff, A new theory of binocular vision, *Arch. Ophthalmol.*, 13, 1935, 151-175.

²⁰ E. A. Bott, Studies on visual flicker and fusion, *Canad. J. Psychol.*, 4, 1950, 145-155.

THE APPARENT ORIENTATION OF A LUMINOUS FIGURE IN DARKNESS

By ERWIN LOUIS HAHN and S. HOWARD BARTLEY,
Michigan State University

Of the many problems of spatial perception, two—shape and orientation—are considered in this study. These problems, though age-old, have in recent years been attacked from several new points of view. Thouless, dealing with the problem of the constancy of shape, has introduced the idea of phenomenal regression, *i.e.* regression to the real object.¹ He differentiated in his study between the *real object* (RO), the *stimulus-object* (SO), and the perceived or *phenomenal object* (PO). This is in essence a way of distinguishing between some physical object or shape that retains its properties regardless of the conditions under which it exists or is perceived by the subject (S) and the actual retinal image of the object received by S. Thouless recognized, however, that the object's apparent shape might be neither that of the real object nor of the retinal image. Other authors, among them Hastorf, have added a fourth object to this list—a conceptual or *assumed object* which we shall designate as AO.² Under some circumstances certain of these objects turn out to be identical, under other conditions one or another of the objects may be missing.

An object's orientation in space is determined in part by its shape, *i.e.* by its SO. In dealing with object-orientation it is advisable, therefore, to use the same experimental conditions as used in the study of shape that the results may be directly compared. Thouless used several sets of conditions ranging from those involving unrestricted cues to those with severely restricted cues. In the study to be reported here, no field-cues whatsoever were present. It was our expectation that with a totally unstructured field—a field without orientational cues—that "regression to the real object" would not occur in the object's phenomenal orientation. Such regression implies some knowledge on S's part aside from the immediate perception.

Although a real object in the objective sense was used the conditions of

* Accepted for publication August 26, 1953.

¹ R. H. Thouless, Phenomenal regression to the real object: I, *Brit. J. Psychol.*, 21, 1931, 339-359; II, *ibid.*, 22, 1931, 1-30.

² A. H. Hastorf, The influence of suggestion on the relationship between stimulus-size and perceived distance, *J. Psychol.*, 29, 1950, 195-217.

the present experiment were so set that no fixed, predetermined, real object (*RO*) known to *S* was exposed to him. It is possible, however, that *S* conceived of the forms presented to him as 'views' of some single object. In that case, instead of a real object (*RO*) in Thouless' sense, he would have an assumed object (*AO*) in Hastorf's sense. Real and assumed objects need not be identical and there need be no single object involved in *S*'s perceptions during the experiments—neither phenomenal (*PO*) nor assumed object (*AO*) need exist. On the other hand, in an experimental such as ours, where repeated presentations are involved, *S* may develop some concept of what is being presented to him and this may come to function as an ingredient in his perception of each and every presentation.

Following Thouless' work, Miller and Bartley used field-glasses to test the effect of optical magnification upon phenomenal regression.³ When the field is structured, as in Thouless' experiments, an object viewed through field-glasses is seen either closer or larger than with the naked eye. In cases of tridimensional objects, there is distortion; the object is flattened in the direction of what Bartley called Chinese perspective.⁴ With plane figures, such as ellipses and circles tilted in the third dimension, it is questionable whether apparent 'distortion'—a change of tilt—is brought about if the field outside the figures is unstructured. It was deduced from Miller and Bartley's results with a structured field that the apparent tilt of the ellipses was reduced, thus paralleling the flattening effect in solid (tridimensional) objects. The evidence was not direct but derived from the fact that phenomenal regression was much less and even absent in some cases when magnification was used. Had the apparent tilts of the objects been determined, the question would have been directly answered.

With the totally unstructured field used in the present study, information regarding object-shape and object-tilt has been obtained. The information regarding tilt or orientation is reported here; that upon shape will be reported in another study.

APPARATUS AND PROCEDURE

Luminous outline figures were presented the *S* on a totally dark field. This was accomplished by using two contiguous dark-rooms; one for the *S* and the other for the exposure-stage and the presentation-equipment. A window, 14 in. high and 28 in. wide, connected the two rooms. Its center was at eye-level when *S* was seated. It could be opened and closed as *E* wished by means of a black draw-curtain.

³ J. W. Miller and S. H. Bartley, Object shape as influenced by instrumental magnification, *J. Gen. Psychol.*, 1954, 50, 141-146.

⁴ Bartley, A study of the flattening effect produced by optical magnification, *Amer. J. Optom.*, 28, 1951, 290-299.

A box, called the stage, was fastened to the wall of the equipment-room opposite the window. Its front was open and larger than the inter-room window. In its back, which was $17\frac{1}{3}$ ft. from *S*'s eyes, there was a 7×7 in. opening in which the panel blocks carrying the test-objects were placed.

All surfaces of both rooms and the stage were painted flat black. The test-objects which were painted with luminescent paint, became visible when the ultraviolet light source—a General Electric CH4 lamp and accessories, located on the stage above the aperture—was turned on.

The test-objects, 12 in number, were made from 12-gauge wire. They were outlines of circles or ellipses and are described in terms of their minor and major axes as follows: 5×5 in.; 4×5 in.; and 3×5 in. These figures were mounted in four different positions with reference to *S*. One of each shape was mounted in the vertical (0°); one was tilted 22.5° from the vertical, away from *S* at the top; one at 45° ; and one at 67.5° . Each figure was mounted on a 7×7 in. panel which fitted in the opening in the back of the stage.

These figures were viewed binocularly by *S* with naked eye and through 2- and 7-power field-glasses. A reduction with an aperture of $2\frac{1}{4} \times 3\frac{1}{4}$ in. was placed 30 in. from *S* when he viewed the figures with naked eyes.

Twenty-one *S*s were used. Their task was to reproduce on a tilt-board the plane in which the figures were seen to lie. The tilt-board, $10 \times 7\frac{1}{2}$ in., pivoted on a central, horizontal axis, was connected to the shaft of a Variac or variable auto-transformer, which lead to a voltmeter in *E*'s room—the readings of which had been calibrated in terms of degrees of tilt from the vertical. Thus the angle at which *S* placed the tilt-board was indicated by the voltage reading. The tilt-board was covered by a white matte card surface and was visible in the dim light supplied by a 7-wa. red bulb suspended overhead.

S's instructions were as follows:

You are to be presented visually a series of forms or figures some of which are mounted in positions other than upright. Your task is to duplicate the plane in which the various figures lie by manipulating the tilt-board in front of you until the plane at which you set the tilt-board best approximates the plane in which you perceive the object to be oriented.

S was allowed all the time he wished to make his judgments but usually he made them promptly. For every condition, distributed randomly throughout a given series, 10 readings were made. When a judgment was returned, it was made known to *E* by means of a buzzer; whereupon he recorded the result, drew the curtain across the window between the two rooms, and placed another test-object in position.

The observations of the first half of the experiment were made with the naked eye; those of the second half were made with field-glasses. In the series with the field-glasses, 12 *S*s used 2-power and 9 used 7-power. The second half of the experiment was performed two days after the first.

RESULTS AND INTERPRETATION

Fig. 1 shows the phenomenal tilts obtained under the three viewing conditions, viz., with naked eyes, with 2-power, and with 7-power glasses.

The solid curves represent the tilts for the three real objects as they assume various stimulus-values or stimulus-shapes. The curve lying to the right is for the circle; the middle curve is for the 4×5 ellipse; and the left-hand curve is the 3×5 ellipse. Thus, if all the responses were to have fallen on

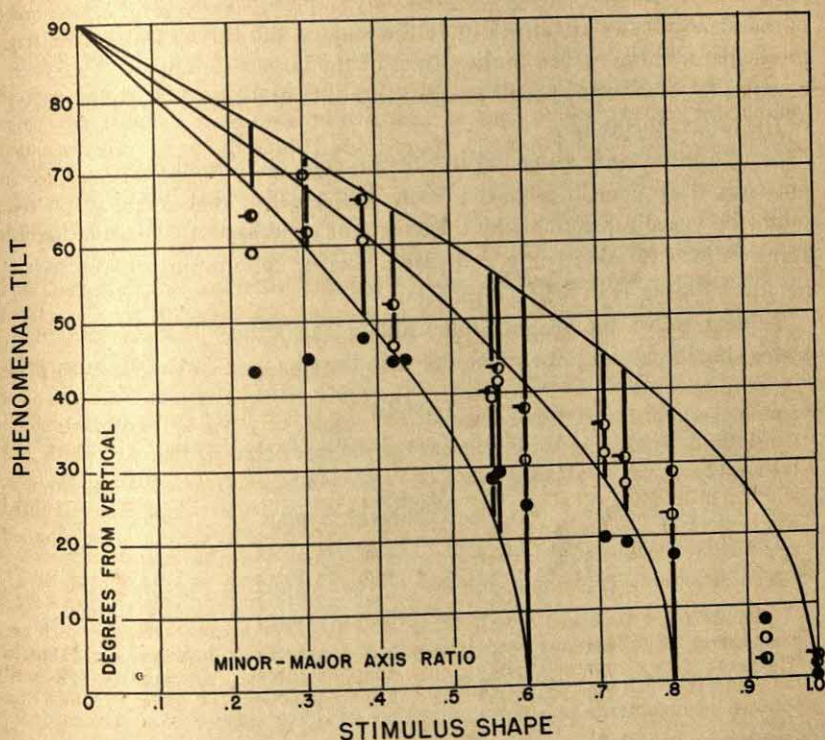


FIG. 1. GRAPH SHOWING BOTH GEOMETRICAL AND PHENOMENAL RELATIONS BETWEEN STIMULUS-OBJECTS AND TILT

The curves show the geometrical relations, and the data depicted in circles, the phenomenal relations. The left-hand solid curve is for the 3×5 ellipse; the middle, the 4×5 ellipse; and the right-hand curve, the circle. The data represent average tilt responses. The solid circles are for responses under 2-power magnification; the open circles for 7-power magnification, and the half-open circles for unaided vision. The vertical bars suggest the range over which the data might likely fall if no determining factors existed in the targets themselves.

or approximately on the right-hand curve, they would have shown that the Ss were responding to the targets as if to tilted circles. Correspondingly, if the responses had followed either of the other two curves, the data would have indicated responses as if to the object the curve represents. As it was, nothing like this happened for any of the targets under

any of the viewing conditions. As would be expected, if no conditions were present to determine the existence of a given real object for the Ss then all tilts within a very wide range would be equally appropriate. This range would extend from verticality for the narrowest target to the appropriate tilt for a circle having the same stimulus-shape. The fact that we did not find any set of responses to follow any of the three curves does not mean that nothing existed in the shapes of the targets to influence the Ss to respond to them as to specific shapes other than the literal stimulus-shapes of the targets themselves.

Since a tilted circle and an ellipse of the same minor-major axis-ratio are not fully equivalent in projected shape, it is possible that this discrepancy might be visually detectable by the Ss in our experiments. If so, it should function as a foreshortening and play a rôle in determining the apparent tilt given the target in the S's responses.

Table I shows the shapes (minor-major axis-ratios) that the observed forms should possess to be consistent with the tilts used in the experiments.

TABLE I

Viewing conditions	RELATION BETWEEN SO AND PO FOR THREE ROs AND THREE VIEWING CONDITIONS					
	RO (5×5)		RO (4×5)		RO (3×5)	
	SO	PO _d	SO	PO _d	SO	PO _d
Unaided eye	.383	4.65×5	.306	3.20×5	.230	2.60×5
	.707	4.33×5	.566	3.87×5	.424	3.40×6
	.924	4.65×5	.739	3.71×5	.554	3.55×5
	1.000	5.00×5	.800	4.50×5	.600	3.80×5
2-power glasses	.383	2.80×5	.306	2.15×5	.230	1.55×5
	.707	3.75×5	.566	3.20×5	.424	2.95×5
	.924	4.65×5	.739	3.90×5	.554	3.10×5
	1.000	5.00×5	.800	4.15×5	.600	3.25×5
7-power glasses	.383	3.95×5	.306	4.40×5	.230	2.15×5
	.707	4.10×5	.566	3.85×5	.424	3.05×5
	.924	4.60×6	.739	4.15×5	.554	3.60×5
	1.000	5.00×5	.800	4.30×5	.600	3.50×5

The results with the three different real objects (5 × 5, 4 × 5, and 3 × 5 minor-major axes) and with the three different conditions of observation (naked eyes, 2-power, and 7-power glasses) are listed separately. The data indicate that the test-objects were responded to differently, not only under the three different conditions of observation but also upon their real stimulus-values. We may conclude that even though there was nothing in the visual field, except the stimulus-objects (targets) themselves, to contribute toward distinguishing between the three real objects, the three were responded to in different ways. The differences must have arisen from the slight dissimilarities between the ellipses and circles tilted to give similar

stimulus-values. This was one of the problems that this investigation was meant to throw some light upon.⁵

The values given under PO_d in Table I might well be called deduced phenomenal shapes to indicate that they are simply the objects which would be calculated from the stimulus-objects (targets) and S's tilt-responses. Consider for example the stimulus-object (SO) having a value of 0.383 (minor-major axis-ratio), the average tilt for the naked eye for the Ss as a group was 65.9° . Now a form, to have such a stimulus-value and such a tilt, must have a minor-major axis-ratio of 4.65×5 . Whether the Ss saw the target as a 4.65×5 ellipse or a form of some other ratio—a circle for example—was not told us. If we choose to entertain some such idea as *regression*, we might say that this value was regressing to a 5×5 shape.

Thouless, in his designation of the real object (RO), to distinguish it simply from the shape as drawn (PO) and the shape (SO) projected on the retina, did not point out that what is to be considered *real* is open to further classification.

The everyday real object is one which we can see, touch, and examine without restriction and still retain the same belief about it. It is independent of the momentary variations due to the unique visual perceptual contact conditions; but this is not the only basis for the real as far as we are concerned. Belief in the nature and properties of a given thing may stem from verbal descriptions or instructions. The concept eventuating in such a case may function with considerable tenacity and form the structuring or organizing principle in subsequent individual perceptual contacts. We might call the object arrived at in this way the assumed object (AO), just as Hastorf 'has done. There is still another 'object' to be reckoned with and it might have been called the phenomenal object had not the term been used as Thouless has done to signify the shape drawn by S when presented with a target. The object to which we have reference is the object perceived in a given situation that may sometimes be different from the drawing made by S. For example, S may see an elliptical target not as an ellipse, in the frontal or near frontal plane but as a tilted circle. Thus the perception is neither a copy of the stimulus-object nor an object arrived at by instruction nor through the influence of the field surrounding the object—*i.e.* through collateral cues. In the situation we are describing, there is something definite about the object itself to effect its being perceived as other

⁵ A further check on this point could be made. A simple test would consist in presenting the two forms, a tilted circle and an ellipse having the same minor-major axis-ratio as the stimulus-value of the circle, side by side a sufficient number of times to obtain the probability of correctness of response—distinguishing between the two.

than the simple stimulus-shape, or, in other cases, as being perceived as nothing but the stimulus-shape. Despite reluctance to add to the already complex classification, we give this perception a name of its own and call it the *observed object* (*OO*).

In the present investigation we are able to say that the targets, although simple, plane, outline forms, possessed enough uniqueness in shape to tend to evoke responses to unique objects (*OO*). These objects (*OO*) and 'Thouless' objects (*RO*) may turn out to be geometrically identical. On the other hand, if the foreshortening effect in tilted circles had not been enough to distinguish them from ellipses of the same shape (*SO*), then the objects (*OO*) would have been the objects presented in the stimulus-shapes (*SO*) of the targets themselves. This might have happened as easily as what was observed in the experiments.

That the objects, *SO*, *RO*, and *OO*, turn out in some cases to be identical is not surprising nor is it a basis for reducing the three categories to a single item in the classification. The identity simply shows how close the organism comes at times to being actually veridical. That this can be so with a simple outline form and in an unstructured field is remarkable.

At any rate we are dealing in our investigation with observed objects (*OO*), not directly with real objects (*RO*). Whether we can speak, therefore, about the observed objects (*OO*) as being those found in the *PO_d*-columns in Table I, or must assume that they are the ideal real objects (*RO*), *i.e.* those known to *E*, is a question. If the latter, then the old problem of 'regression' toward such objects returns because the phenomenal objects (*PO*) obtained in the experiments were only approaches to the values 5×5 , 4×5 , and 3×5 of the real objects. To answer our question, we must have additional information, including protocols from our *Ss* regarding the shapes that they were consciously seeing. Such protocols would not, however, give complete and final answers in all respects, for we have already found some features of the *Ss*' responses that function on a level below verbalization. From these responses one would have thought that the *Ss* were well aware of certain characteristics of the targets, and yet, under questioning, no evidence was obtained to show this was so.

The results for the 7-power glasses were unexpected. They were more like those for the unaided eye than were the results with the 2-power glasses. There seems to be no way of accounting for this except to suppose that, owing to the short viewing distance, the 7-power glasses brought the objects so close to the viewer as to introduce some 'unreal' effect, and thus to distort the results. This sometimes happens in perceptual experiments, in which case some variable that is producing regular shift in results sud-

denly fails to be a factor and the organism responds much as if the variable now possessed a small or zero value. This is the case in the use of size-lens for increasing percentages of meridional magnification. Our suggestion, therefore, is that to attempt to account for the unexpected results with the 7-power glasses, a series of ascending powers between 2- and 7-power glasses be used. Such a procedure might well reveal a regular effect up to a certain point and from there on, a reversal.

All of the differences that have been mentioned so far are significant beyond the 1-% level of confidence. Actually, there was also a significant difference between the ways in which the *Ss* themselves behaved. This is to be expected for the situations were so nearly unstructured as to allow for any one of a number of outcomes. That every conceivable outcome was not exemplified in the results was possibly due, in part, to not having a vast number of *Ss*, and also was due to some outcomes being much more probable than others. The central issue in the study is not the demonstration of how great an array of varying responses will eventuate, but a determination of the character of the interrelationships between various aspects of the perception of forms. The investigation is incidentally a demonstration that the nature of perception depends upon the task involved, that is, what is asked of *S* as well as upon the geometry of the targets.

It would seem that even when the field outside the figures was unstructured, the tilted figures tended to be treated as such. All of the *Ss* used in this investigation were first used in obtaining shape responses (drawings) to the same targets. In that investigation, nothing was said about tilt or target orientation. It is conceivable that the behavior with reference to the targets not only depends upon what the immediate task is but upon previous tasks with reference to the same targets. This factor was not studied.

SUMMARY

The apparent tilts of a number of plane outline forms (circles and ellipses) variously oriented with reference to the *S* were obtained. These forms were made luminous with ultraviolet light and thus were the only visible structure in the fields. They were viewed under three conditions; namely, with the unaided eye, under 2-power magnification by field glasses, and 7-power magnification.

Despite the total absence of field-structure outside the targets they were readily seen as tilted. Circles and ellipses tended to be distinguished.

This was made barely possible through the slight difference in shape between ellipses and circles that were tilted so as to have the same minor-major axis-ratios. This is what is ordinarily known as foreshortening.

Since *Ss* were unaware of the 'true' nature of the targets, Thouless' real object (*RO*) was not directly applicable. Instead, Hastorf's assumed object (*AO*) possibly played a rôle. To further clarify the matter of response in general, an additional item was added to the classification of objects. This was called the observed object (*OO*), the object accruing from conditions within the target itself without the participation of differentiations in the field conventionally called cues. The doctrine of cues was seen to be inapplicable but a discussion of this was omitted. Such a discussion will appear in a later paper from this laboratory.

FATIGUE IN TWO SIMPLE VISUAL TASKS

By CURT BERGER and AXEL MAHNEKE, University of Copenhagen

This research was prompted by the difficulty which often is encountered in attempts to find quantitative and statistically significant evidence of deterioration in the course of prolonged visual work, despite subjective reports of 'tiredness.'¹ Its purpose was to describe the development of fatigue, measured in terms of decrement in performance, during the exercise of two relatively circumscribed visual functions.

Experiment I: Landolt's Ring. With his head held in position by a chin-rest (but permitted to move his eyes and to blink), *O* sat before a white background subtending a visual angle of 10° and viewed a broken ring monocularly through an artificial pupil 2 mm. in diameter. The apparatus, previously described,² permitted automatic variation of the distance between *O* and the test-object at a constant velocity of 5 cm. per sec. To register his reports, *O* pushed a button which operated an electromagnetic perforator. Retinal illumination was held constant (at 3.5 lux) with a Sorensen voltage-regulator and was controlled by a luxmeter. For a period of 50 or 60 min., *O* made a continuous series of acuity-determinations, following which there was a 5-min. period of rest in the dark. Then a series of 20 additional determinations followed.

The results of three *O*s are shown in Fig. 1, each point there plotted being the mean of 10 determinations. As the curves show, visual acuity decreased in the course of work for all *O*s (18–32%), although the variability of the settings (expressed in terms of the *SD* of the settings) showed no systematic change. Acuity increased after the 5-min. rest, but it usually did not attain the pre-work level. At the conclusion of the fatiguing period—before the resting interval—all *O*s reported feelings of tiredness. Two suffered from headaches, located above the observing eye, which lasted from 30 to 60 min. after the experimental hour.

Experiment II: Critical fusion frequency. Decrease in the critical fusion frequency (*CFF*) of flicker has often been used as an index of central fatigue on the assumption that the adequacy of neural functioning is directly proportional to the maximal impulse-frequency which can be differentiated.³ More than 20 years ago,

* Accepted for publication May 30, 1953. This study, from the Tscherning Laboratory, was supported in part by a grant from the Danish Government.

¹ Leonard Carmichael and Walter Dearborn, *Reading and Visual Fatigue*, 1947, 1-473. The relation between feelings of tiredness and level of performance is, of course, an important problem (Madison Bentley, Is fatigue tiring? this JOURNAL, 57, 1944, 264 f.)

² Curt Berger, Some experiments on the width of symbols as determinants of legibility, *Acta Ophthalmol.*, 26, 1948, 4 f.

³ Ernst Simonson and Norbert Enzer, Measurement of fusion frequency of flicker as a test for fatigue of the central nervous system, *J. Indus. Hyg. Tox.*, 23, 1941, 83-89.

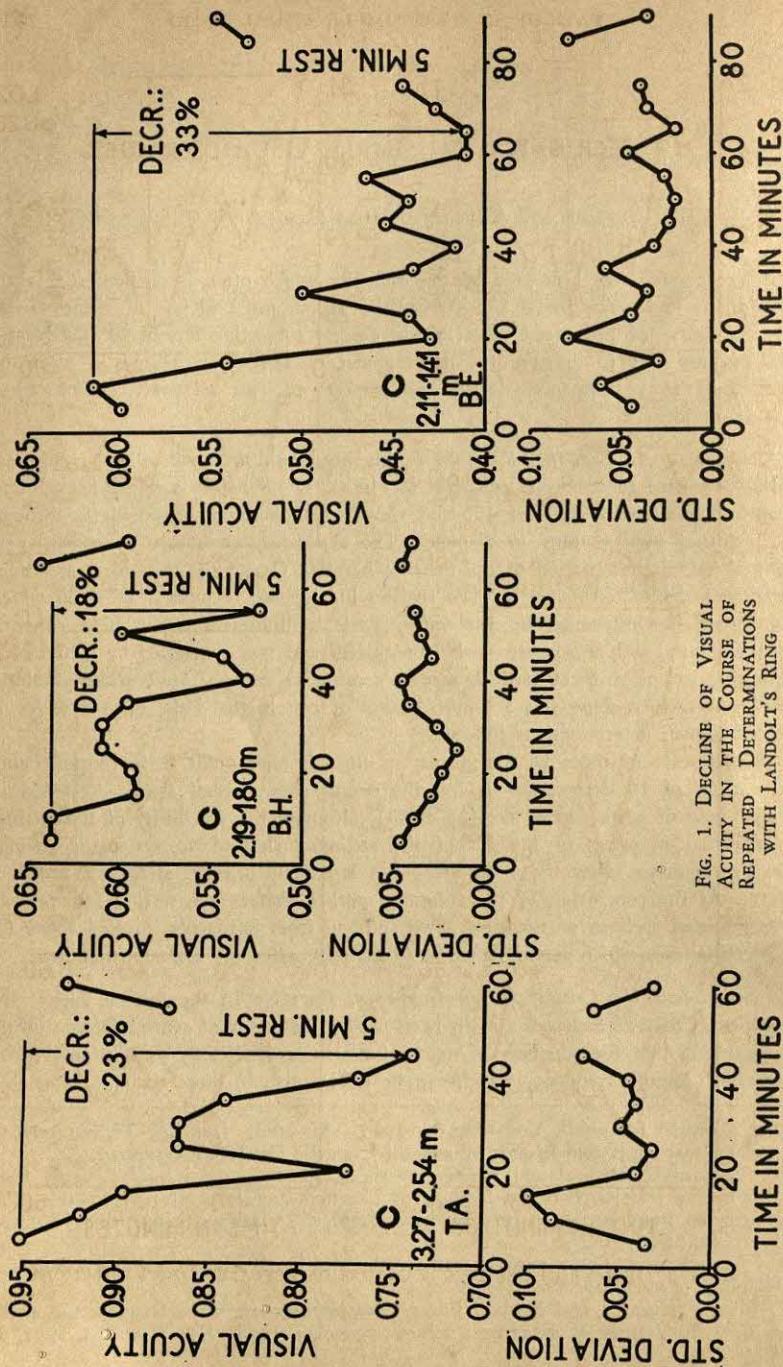


FIG. 1. DECLINE OF VISUAL ACUITY IN THE COURSE OF REPEATED DETERMINATIONS WITH LANDOLT'S RING

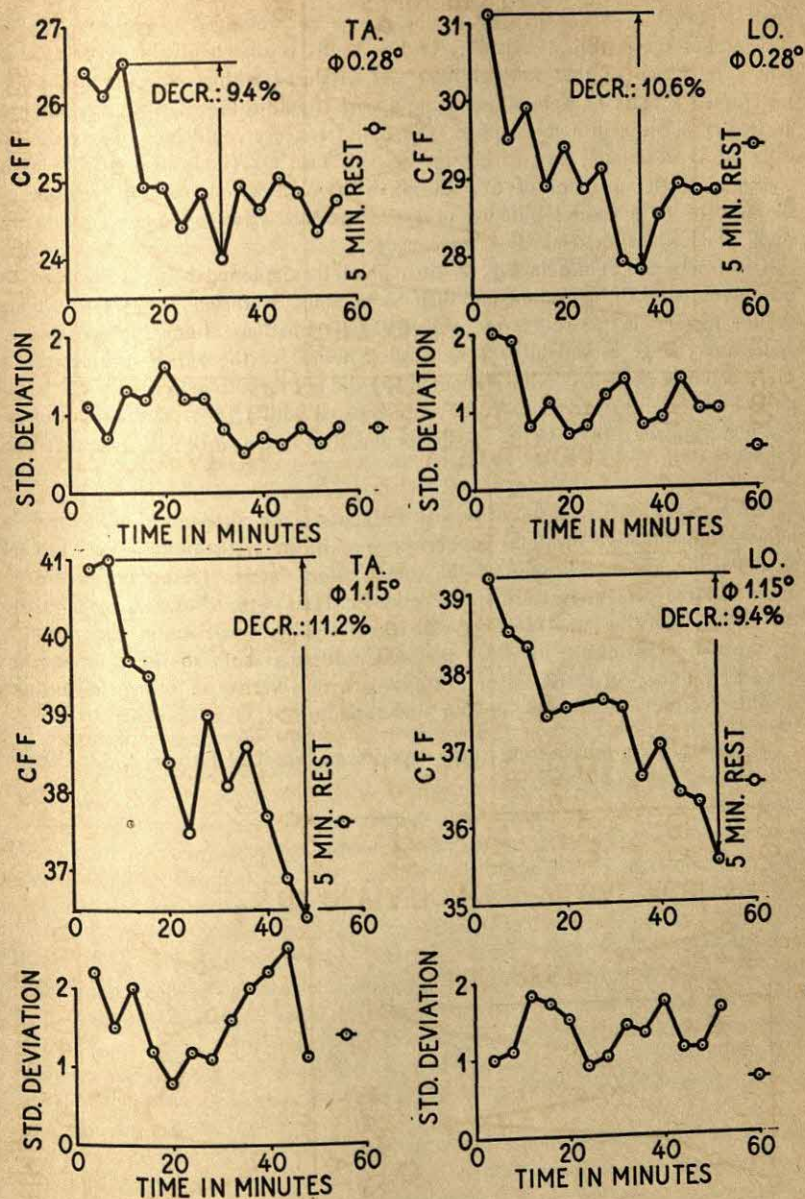


FIG. 2. DECLINE IN CFF IN THE COURSE OF REPEATED DETERMINATIONS

Snell reported a decline in *CFF* in the course of prolonged exposure to visible flicker,⁴ but the statistical reliability of his results is sufficiently in doubt to warrant a repetition of that work. Flicker was produced with a rotating disk which interrupted a beam of light illuminating a perfectly-diffusing glass surface through a circular opening in a metal plate. The flash-frequency could be varied automatically by approximately two flashes per second. Two *O*s (both emmetropes) made continuous determinations of *CFF* during a 50-min. period, with no attempt on the part of *E* to control blinking or eye-movements. After a 10-min. rest in the dark, 20 additional determinations were made.

The results are shown in Fig. 2, each point there plotted being a mean of 10 measurements, half ascending and half descending. Both *O*s showed a striking decline in *CFF* in the course of continued determinations although there was no systematic change in variability (expressed in terms of the *SD* of measurement.) The amount of percental decline (about 10%) did not seem to be related to the angular size of the flickering target (0.28° as compared with 1.15°). The amount of recovery during rest (13–15%) seemed to be less than in the work with Landolt's ring (70–82%).

Interpretation. The course of visual fatigue can be determined accurately in terms of quantitative measures of performance if the test-stimulus itself is used as the tiring agent and some of the compensatory factors (rest-periods) are reduced. In the present instance quite delimited tasks were chosen for study, and the testing situation was identical with the fatiguing situation on the principle of muscular experiments, in which the weight lifted is both the measure and the cause of fatigue. An extension of this approach to a variety of other simple tasks should contribute significantly to our understanding of visual fatigue.

⁴ P. A. Snell, An introduction to the experimental study of fatigue, *J. Soc. Motion Pict. Eng.*, 20, 1933, 367-390.

BINOCULAR RIVALRY OF ACHROMATIC COLORS

By HANS WALLACH and PAULINE AUSTIN ADAMS, Swarthmore College

In an earlier paper one of the present authors reported experiments which strongly suggest that achromatic surface-colors (white, gray, black) are the result of a process of interaction.¹ This is certainly true of the surface-colors that are perceived in a dark room when adjacent areas of the retina are stimulated by two different intensities of neutral light. Not only are the achromatic colors seen under these circumstances dependent on the relation of the two light-intensities rather than on the absolute intensities, but stimulation by two light-intensities also represents the simplest condition under which surface-colors are seen; a single bright area in a dark field does not look white or gray but has a luminous appearance. Only upon the introduction of a second bright area of different intensity (adjacent with the first and preferably surrounding it) will the luminous appearance change into that of a surface-color, which will be white, gray, or black, depending on the relation of the light-intensities in the two areas. If the intensity of the surrounding area is one-half the intensity of the inner area, the inner area will look white; if the intensity of the outer area is twice that of the inner area, the inner area will look light gray; as the intensity of the outer area is further raised, the gray of the inner area becomes darker and darker, with a good black appearing at a ratio of approximately 20:1.

In the light of these results it is interesting to examine the retinal rivalry which occurs when different achromatic colors are presented to corresponding regions of the two eyes. Does rivalry depend on the difference in stimulation or on the result of interaction, *i.e.* on the difference in the achromatic colors that would be seen if the left-eye-arrangement or the right-eye-arrangement were presented alone?

The usual procedure of demonstrating retinal rivalry is compatible with either interpretation. When, in a stereoscope, a white figure is presented to one eye and a congruent black figure to the other eye, both on a gray ground and so arranged that their images fall on corresponding retinal regions, strong rivalry is experienced. In the retinal images of the two figures, the intensity of stimulation under these conditions differs by a factor of twenty, and, by viewing monocularly, *O* would see a black figure with one eye and a white figure with the other. It is possible, however, to present neutral light stimuli in such a way that achromatic surface-colors will not be perceived. It also is possible to make *O* see different achromatic colors in each monocular view when the intensities of stimulation in the corresponding areas are the same. The usual coincidence of difference in color with difference in stimulation can thus be avoided.

Method. The experiments were done with an arrangement similar to that described in the previous paper. Of two identical lanterns placed side by side in a dark room, one projected a circular area of light (disk) on a screen and the other projected a

* Accepted for publication June 29, 1953.

¹ Hans Wallach, Brightness constancy and the nature of achromatic colors, *J. Exper. Psychol.*, 38, 1948, 310-324.

ring which surrounded the disk, the inner contour of the ring coinciding with the edge of the disk. Another pair of lanterns produced an identical ring-disk arrangement at some distance from the first. The diameter of the disks was 6.50 in. and the width of the rings 4.25 in.; the distance between the centers of the disks was 50 in. By a procedure described in the previous paper, the lights of the lantern beams were so adjusted at the outset that all four areas were of the same intensity. The intensity of each area could be altered independently in well-defined steps by placing episcotisters with openings of different size in the lantern beams.

The *O* was placed at a distance of 14 ft. from the screen in a symmetrical position with relation to the two ring-disk arrangements. He had before him a set of four first surface mirrors comprising a stereoscope which made it possible for him to see a different ring-disk arrangement with each eye. When the mirrors were properly adjusted, a fused image of the two ring-disk arrangements appeared straight ahead of *O* at the distance of the screen.

Experiment 1. In the first experiment the intensities were so adjusted that the colors of the two disks were different enough to produce rivalry when the two ring-disk arrangements were fused. For the left eye, the ring was at full intensity and the disk at half that intensity (the episcotister in the beam of the disk was opened 180°); under these conditions the disk was light gray. For the right eye, the ring was at full intensity, but the disk was at only $1/12$ of the maximal intensity (30° episcotister-opening); in direct view, therefore, the disk appeared dark gray. When the lanterns which projected the rings were turned off and only the disks were visible, the disks appeared as luminous spots, the left one bright and the right one rather dim.

Ten *Os*, undergraduates of Swarthmore College, served in this experiment. All had previous experience with binocular rivalry, having served weeks before in another experiment. They were presented (individually) first with the ring and disk arrangements described above. When the two patterns were fused stereoscopically, all 10 *Os* reported rivalry. The period of observation was limited to 10 sec. to guard against the appearance of after-images which might be reported as rivalry. Binocular rivalry between the light and a dark gray consisted in a slow but steady sequence of changes between lighter and darker colors; careful observation was necessary to notice it. The rings were then eliminated and the *Os* stereoscopically fused the resulting luminous spots. Nine of the 10 *Os* did not detect any rivalry although they were looking for it. Only one *O* reported rivalry, and we suspect that here a sudden eye-movement caused him to mistake an after-image shifted from the disk-area for rivalry.

These results indicate that differences in stimulation will lead to rivalry only under conditions in which *surface-colors* are perceived. The experiment demonstrates the rôle of the interaction-process in a rather striking fashion. Since rivalry occurred only when the rings were added, and since the two rings were of the same intensity, the cause of rivalry must lie in the juxtaposition of ring and disk. Only the difference in the *relation* between the intensities of ring and disk can have been responsible.

Experiment 2. This experiment was similar to the first except that the colors produced by interaction were white and gray instead of light gray and dark gray. A

group of 10 Os different from those of the first experiment were presented first with the luminous disks alone and then with ring-disk arrangements yielding surface-colors. The disk-intensities were again produced with episcotister-openings of 180° and 30° , and when they were presented without rings the disks appeared luminous, one bright, the other dim. For each ring, the episcotister-opening was 90° . When the rings were present they made the more intense disk appear white and the other one gray. All 10 Os reported fusion when the luminous disks were stereoscopically combined, but they reported strong rivalry when the rings were added. These results confirm those of Experiment 1.

It should be mentioned that observation of rivalry in this experiment is complicated by the fact that the contours of the white and of the gray disks do not fuse. It is only by virtue of the fusion of the outer ring-contours that the two monocular patterns stay stereoscopically combined. The difficulty stems from the fact that opposite gradients give rise to the formation of the inner contours in the two monocular arrangements.

Experiment 3. This experiment was designed to demonstrate in the most direct fashion that rivalry depends on interaction, *i.e.* on the difference in the achromatic colors as seen in each monocular view rather than on the difference in stimulation as such. The hypothesis was tested that a difference in stimulation which does not lead to a difference in perceived color fails to produce rivalry.

TABLE I
RESULTS OF EXPERIMENT III

Test	Episcotister-opening (degrees)				No. of Os reporting	
	left eye		right eye		fusion	rivalry
	ring	disk	ring	disk		
1	90	60	360	40	0	24
2	90	60	360	240	22	2
3	90	10	360	40	20	4
4	90	60	90	40	24	0
5	90	—	360	—	24	0
6	—	60	—	40	24	0

To produce rivalry with approximately equal intensities in the two disks, a lantern beam passing through an episcotister with a 60° -opening produced the disk for the left eye and a beam through a 40° -opening produced a disk for the right eye. Markedly different colors were given the disks by surrounding the left one with a ring whose light passed through an episcotister-opening of 90° while the ring of the right disk was of maximal intensity. The ring-disk ratio for the left eye was 3:2 (which gave a very light gray disk), and the ratio for the right eye was 9:1 (very dark gray disk). These settings were used for Test 1 (Table I). The tests in which the ring-disk ratios were the same for both eyes were made by keeping intact the Test-1 arrangement either for the left eye or the right eye and varying the other appropriately (Tests 2 and 3 of Table I). Test 4 was a control to demonstrate that disks at episcotister-openings of 40° and 60° surrounded by rings of equal intensity would not give rise to rivalry. Test 5 was intended to demonstrate that rings alone

would not produce rivalry despite marked differences in intensity. Test 6, as shown in the table, completed the experimental design. For this experiment, 24 inexperienced *O*s (undergraduates of Swarthmore College) were used. They were first made familiar with the phenomenon of rivalry by letting them view in a mirror-stereoscope a light gray disk with one eye and a black disk with the other eye, both on a medium gray background. Following this training, each *O* was given the six tests in the order: 5, 2, 1, 3, 6, 4.

The results are shown in the last column of Table I. Only Test 1 produced consistent reports of rivalry, whereas Test 2 and Test 3, as well as the other tests, produced a strong preponderance of reports of fusion. All results are significant at better than the 1-% level of confidence, and we are inclined to attribute the six rivalry reports obtained in Tests 2 and 3 to shifting after-images. This experiment demonstrates, therefore, that when disks of nearly identical intensity appear as different colors (Test 1), rivalry develops consistently, but, when disks of markedly different intensity appear almost identical (Tests 2 and 3), fusion is successful.

Conclusion. Interaction-conditions determine the mode of appearance of neutral colors and the brightness of achromatic surface-colors. Whether two areas stereoscopically combined will display binocular rivalry also depends on interaction rather than on local stimulation.

AN ANALYSIS OF THE FORMATION OF VISUAL DISCRIMINATIVE HABITS OF RHESUS MONKEYS

By J. M. WARREN, University of Texas

An early variant of S-R theory states that the difficulty of a discriminative problem involving recurrent stimuli is a function of the differential excitatory strengths associated with the positive and negative stimuli; the excitatory strength of a stimulus being dependent on the relative number of times it has been reinforced or non-reinforced in previous discriminations.¹ Thus, if the positive member of a new pair has been rewarded more frequently than the negative in previous pairings with other stimuli, this new problem will be solved quickly, and the number of errors made by individual Ss will be correlated (negatively) with the magnitude of the difference in frequency of reinforcement in favor of the positive stimulus.

Spence obtained significant rank order correlations between *R* (a measure of the difference in the excitatory strengths of the positive and negative stimuli) and performance on a series of tests involving recombinations of a small number of stimuli. He concluded that "discrimination-learning is a cumulative process resulting from differential reinforcement and non-reinforcement of responses to relevant stimulus-components."²

Lashley later tested Spence's theory with respect to the relevance of stimulus-components and the cumulative effects of reinforcement, and demonstrated that differential reinforcement affects only those components of a compound stimulus toward which *S* is set to respond.³ Since this demonstration of perceptual dominance in 1942, S-R theorists have largely restricted themselves to consideration of stimuli varying in a single dimension. Lashley also repeated Spence's experiment in an attempt to evaluate the adequacy of cumulative reinforcement and non-reinforcement as a sufficient explanation of discriminative learning. He found, however, that correlational analyses, based on small samples, were not sufficiently precise to distinguish the alternative predictions of continuity and non-continuity theory.

More recently, studies of the formation of learning sets in primates suggest that Spence's experimental design may yet provide a critical test of continuity theory.⁴ Especially relevant is Riopelle's observation that both positive and negative transfer are suppressed as the formation of learning set progresses: "During the first phase

* Accepted for publication August 5, 1953. This work was conducted at the Radiobiological Laboratory of The University of Texas and United States Air Force, and was supported (in part) by funds provided under Contract AF 18 (600)-165 with the USAF School of Aviation Medicine, Randolph Field, Texas.

¹ K. W. Spence, Analysis of the formation of visual discrimination habits in chimpanzee, *J. Comp. Psychol.*, 23, 1937, 77-100.

² Spence, *ibid.*, 97.

³ K. S. Lashley, An examination of the "continuity theory" as applied to discriminative learning, *J. Gen. Psychol.*, 26, 1942, 241-265.

⁴ H. F. Harlow, The formation of learning sets, *Psychol. Rev.*, 56, 1949, 51-65.

of the experiment performance on reversal problems was negatively correlated with performance on the original problems, but the correlation approaches zero with increased practice."⁵

Spence used chimpanzees which had quite variable amounts of training on visual discriminative tasks, and it is impossible to estimate the degree of learning set attained by the experienced Ss. In view of Riopelle's findings regarding the suppression of transfer during the formation of learning set, it is predicted that a group of relatively sophisticated monkeys would show little or no transfer within a series of tests such as Spence used, and thus yield very low order correlations between response-strength and performance. All the Ss which served in the present study had previously received 75 days' training on visual discriminative tasks. This homogeneous group provides an opportunity for comparing the prediction of Spence's theory with the predictions implicit in the studies of learning set.

METHOD AND PROCEDURE

Subjects. The animals (12 rhesus monkeys) used as Ss in this study had identical testing experiences: 15 days' training on a 15-pair concurrent discrimination habit; 36 days' on visual acuity problem; 5 days' retraining on the concurrent discrimination; and 18 days' testing on a series of object-discriminations.⁶

Apparatus. The Wisconsin General Test Apparatus was used throughout.⁷ The stimuli were 3 sq. in. black figures (cross, circle, triangle, and rectangle) cut from 3/8-in. plywood and nailed on 3 by 3 in. white squares. The identical background square were 5/8-in. pine stock; similar objects were illustrated in a publication by Harlow.⁸

Procedure. Each of the six paired combinations of the four stimulus-objects was presented to 2 of the 12 Ss each day, the order of presentation being counterbalanced by use of a latin square. (The same latin square was replicated, since it was uncertain that laboratory requirements would permit all Ss to complete testing.)

Thus, the six paired comparisons of the forms were orthogonal to ordinal position and the possibility of interaction between differential difficulty of the several comparisons arising from stimulus preferences and the pattern of reinforcement was minimized.

The designation of positive and negative stimuli shown in Table I is essentially identical with Spence's through Test 2; Test 3 was substituted for Spence's test of the same number, because results in agreement with continuity theory would most probably be anticipated on this test since the now positive D-stimulus had been incorrect on two previous occasions. The now negative A-stimulus had been correct in two preceding problems. Test 4 is not strictly comparable with any of Spence's tests. One problem, consisting of 100 non-correction trials, was presented each day for six testing days.

⁵ A. J. Riopelle, Transfer suppression and learning sets, *J. Comp. & Physiol. Psychol.*, 46, 1953, 113.

⁶ J. M. Warren, Some stimulus variables affecting the discrimination of objects by monkeys, *J. Genet. Psychol.* (in press).

⁷ Harlow, *op. cit.*

⁸ Harlow, Studies in discrimination learning by monkeys: V. Initial performance by experimentally naïve monkeys on stimulus-object and pattern discrimination. *J. Gen. Psychol.*, 33, 1945, 3-10.

RESULTS

Table I shows the number of animals which attained the criterion of 16 correct responses in the last 20 trials ($P < .01$), and the mean percentage errors in 100 trials, on the six discriminative problems. Continuity theory implies that the preliminary habits and Test 2 should be approximately equal in difficulty since neither stimulus has acquired excitatory strength through reinforcement. The theory further predicts that Tests 1, 3, and 4 should be more difficult than the preliminary habits and Test 2, since the negative stimulus is either approximately equal (Test 1), or

TABLE I
PERFORMANCE ON THE SERIES OF DISCRIMINATIVE PROBLEMS

Problems	Stimuli		% errors (all Ss)	Number attaining criterion	% errors
	positive	negative			
Preliminary: Habit 1	A	B	16.4	12	16.4
Preliminary: Habit 2	C	D	20.3	12	20.3
Test 1	A	C	18.3	10	19.3
Test 2	B	D	25.8	9	18.7
Test 3	D	A	54.7	4	43.0
Test 4	C	B	23.0	9	13.2

greater (Tests 3 and 4), in excitatory strength than the positive. Analysis of variance reveals that variation due to ordinal position in the series of discriminations (*i.e.* pattern of reinforcement) was the only significant variable, but Tukey's gap-test indicates that only Test 3 differs reliably from the mean of the remaining tests and the preliminary habits.⁹ Thus, these results confirm continuity theory in only one of three cases.

The cumulative effects of reinforcement and non-reinforcement in the preliminary habits and Tests 1 and 2 on performance in the subsequent tests are summarized in Table II, which gives the correlation of *R* (defined by Spence as a measure of the

TABLE II
RANK ORDER CORRELATION BETWEEN *R* (EXCITATORY STRENGTH)
AND TWO MEASURES OF DISCRIMINATIVE LEARNING

Test	Range of <i>R</i> -values	Correct responses: Trials 1 to 20	Total errors
1	72	-.53	+.41
2	72	-.31	+.44
3a	58	-.17	-.23
3b	84	+.12	-.35
4a	35	+.06	+.06
4b	59	-.21	-.15

relative strengths of the positive and negative stimuli, determined by the number of reinforcements they had previously received) with the number of correct responses in the first 20 trials, and with the total number of errors. The values of *R* for

⁹ J. W. Tukey, Comparing individual means in the analysis of variance, *Biometrics*, 5, 1949, 99-114.

Tests 1 and 2 were calculated according to the same formulae used by Spence (for Test 1: R = the number of times A was reinforced minus the number of times C was reinforced. For Test 2: R = the algebraic difference of the number of times D was non-reinforced minus the number of non-reinforcements received by B). Two methods were used in computing the value of R for Tests 3 and 4 based on the two possible interpretations of Spence's theory. In Case *a*, it was assumed that the positive stimulus had zero response strength, since it had been extinguished (Test 4) or had never been reinforced (Test 3); the value of R is then simply the negative of the number of times that A or B had previously been rewarded. In Case *b*, the now positive stimuli were assumed to have negative excitatory strength equal to the number of previous non-reinforcements, and R was, therefore, equal to the algebraic difference of the number of times the positive stimulus had been non-reinforced minus the number of times the negative stimulus had been reinforced. Since three S s failed to attain the criterion of learning on Tests 1 or 2, the correlations for Tests 3 and 4 are based on nine S s only. It is apparent that only low and inconsistent correlations were obtained; the single significant value, ($\rho > .51$), the correlation between R and correct responses on Trials 1 to 20 of Test 1, is opposite in sign to the comparable correlation reported by Spence.

Thus, the data of this study indicate that Spence's theoretical schema for discriminative learning does not hold for even partially sophisticated S s. It is true that a significant increase in errors occurred on Test 3 and this finding indicates that reinforcement and non-reinforcement of individual stimuli has some effect on learning to discriminate between them. The failure to obtain significant correlations between R and performance, and the fact that Tests 1 and 3 were no more difficult than theoretically easier discriminations, suggest, however, that the predominant response tendency for these monkeys was to approach each recombination of stimuli as new problem, in spite of discrepancies in the number of times the positive and negative stimuli had been rewarded. This interpretation is supported by the observation that the S s manifested no strong preferences for any stimulus.

SUMMARY

This experiment, a partial repetition of Spence's study with chimpanzees on the effect of antecedent reinforcement and non-reinforcement on a series of discriminative learning problems, was performed in order to test the generality of Spence's results and the continuity theory which predicted them. The result of one of three critical tests was in agreement with the implications of continuity theory, but the only correlation between response strength and performance which was significant at even the 5 per cent level of confidence indicated the inverse of the relationship predicted by continuity theory. The results suggest that continuity theory can not, without serious modification predict the performance of test-wise primates.

COMPARISONS OF INCIDENTAL AND INTENTIONAL LEARNING AFTER DIFFERENT NUMBERS OF STIMULUS PRESENTATIONS

By IRVING J. SALTZMAN and RITA L. ATKINSON, Indiana University

In a recent paper by Neimark and Saltzman it was found that the difference between the scores of incidental and intentional learning depends on the rate of presentation of the learning material.¹ After four presentations of a list of 14 numbers where each number appeared for 2 sec., no difference was found between the scores of incidental and intentional learning. When, however, the numbers were presented at a slower rate of 6 sec. per number, the scores of intentional learning were significantly higher than those of incidental learning. While the scores of the incidental learners did not change significantly from the fast rate to the slow rate, the scores of the intentional learners, on the other hand, increased significantly.

A similar relationship might be expected when the length of the time of exposure is held constant, but the number of exposures per item is increased. Specifically, one might expect that after only a few presentations of a list of numbers, no difference would be found between the scores of intentional and incidental learning, but that after many presentations of the list, a difference would be found. The present study was designed to investigate this relationship between the number of trials, on the one hand, and the difference between incidental and intentional learning scores, on the other.

Procedure. The Ss were 160 Indiana University undergraduate students randomly divided into two groups of 80 each—an incidental learning group and an intentional learning group. Each group was further divided into four subgroups of 20 Ss each. The learning material was a list of 14 two-digit numbers selected from a table of random numbers. A different number of presentations of the list was given each of the subgroups. The subgroups in each group received either 2, 6, 8, or 16 presentations of the list. The rate of presentation of the list of numbers was the same for all subgroups: each number appeared for 3 sec. and then was replaced by the next number. The numbers were presented by means of an electrically operated memory-drum. A 10-sec. period intervened between successive presentations of the list. Four different random orders of the 14 numbers were used.

The Ss in the incidental learning group were read the following instructions.

We are interested in finding out the best way to code numbers. Some numbers will appear one by one in this slot. As soon as a number appears, we want you to circle that number on the coding matrix. The numbers range from 11 to 99. After all the numbers have appeared, there will be a short rest-period. Then we will repeat the procedure with a different coding matrix. We have (2, several) different

* Accepted for publication September 3, 1953.

¹ Edith Neimark and I. J. Saltzman, Comparisons of incidental and intentional learning under different rates of stimulus-presentation, this JOURNAL, 66, 1953, 618-621.

kinds of matrices, so we will repeat the procedure (2, 6, 8, 16) times. During the rest-period you will be able to inspect the next matrix to see how the numbers are set up.

The matrices were mimeographed sheets containing the numbers 11 to 99 arranged in sequence. On one matrix, (a), the numbers were arranged in four columns, on another, (b), in five columns, on the third, (c), in nine columns, and on the fourth, (d), in 10 columns. The numbers were arranged in sequence vertically on the matrices (a) and (c) and horizontally on matrices (b) and (d).

The Ss in the intentional learning group were read the following instructions.

We want to see how fast you can learn numbers. There will be 14 numbers ranging from 11 to 99. When a number appears in this slot, you are to circle that number on this sheet. A short rest-period will follow each presentation of the numbers. There will be (2, 6, 8, 16) presentations. During the rest-period you will be given another sheet with the numbers arranged somewhat differently. You are to repeat the procedure with the new sheet. Remember, you are to learn as many of the numbers as you can.

For the two subgroups where the list was presented only two times, half of the Ss in each subgroup used matrices (a) and (b), and the other half used matrices (c)

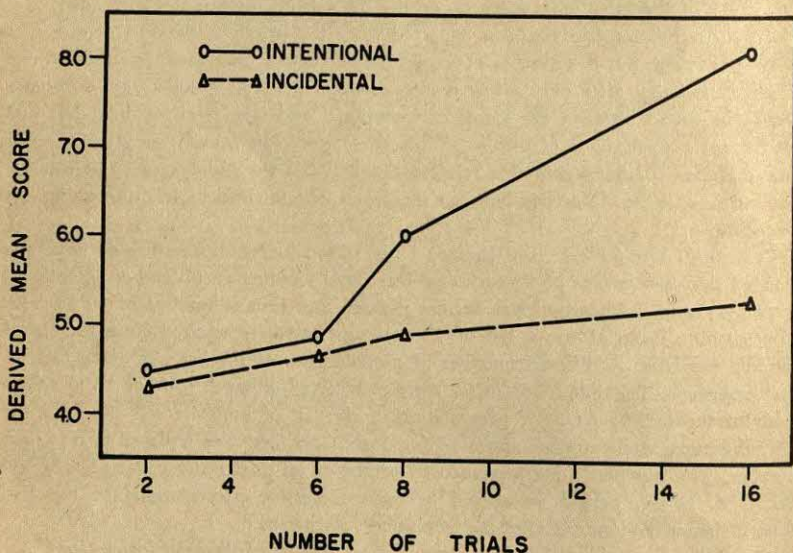


FIG. 1. DERIVED SCORES FOR THE INTENTIONAL AND INCIDENTAL LEARNING GROUPS

and (d). For the two subgroups where the list was presented six times, half of the Ss in each subgroup used matrices (a) and (b) twice and matrices (c) and (d) once, and the other half used matrices (c) and (d) twice and matrices (a) and (b) once. For the subgroups where the list was presented eight times and 16 times, each S used all four matrices twice and four times, respectively.

Immediately following the last presentation of the list of numbers, each S was

given a recognition-test. The 14 numbers of the learning series plus 42 additional two-digit numbers were randomly arranged in four columns on a mimeographed sheet. Each *S* was requested to circle all the numbers which he recognized as having appeared on the memory drum. Two minutes were allowed for the test.

Results. Scores on the recognition-test were derived by subtracting from the number of correctly circled numbers, the number that might be expected to be correct by chance. One of every four numbers circled was considered to be correct by chance. For example, if *S* circled 8 numbers, 7 of which were correctly circled, his score would be 7 minus 2 ($8/4 = 2$), or 5. Means of the derived scores were computed for each subgroup and are shown graphically in Fig. 1. An analysis of variance yielded an *F* of 12.90 for the difference between the intentional and incidental learning groups and an *F* of 11.90 for the differences between the different numbers of presentations of the list. Both of these values are significant at the 0.1-% level of confidence. Table I shows the means and the differences between the means of the

TABLE I
MEANS AND DIFFERENCES BETWEEN MEANS OF DERIVED SCORES

Number of presentations	Groups		Diff.	t	p
	Intentional learning	Incidental learning			
2	4.46	4.28	0.18	0.50	.616
6	4.85	4.66	0.19	0.28	.764
8	5.98	4.89	1.09	1.85	.064
16	8.12	5.31	2.81	4.44	.001

intentional and incidental learning subgroups for each of the different numbers of presentations of the list. The *t*-values for the differences and their corresponding probabilities are also shown. Table II contains comparisons of the mean scores of

TABLE II
DIFFERENCES BETWEEN THE DERIVED MEANS OF THE SUBGROUPS IN EACH GROUP

Subgroups compared	Groups			Groups		
	Intentional learning			Incidental learning		
	Diff.	t	p	Diff.	t	p
6-2	0.39	0.66	.484	0.38	0.74	.482
8-2	1.52	3.38	.001	0.61	1.08	.270
16-2	3.66	7.14	.001	1.03	1.77	.072
8-6	1.13	1.71	.088	0.23	0.36	.688
16-6	3.27	4.66	.001	0.65	1.07	.272
16-8	2.14	3.61	.001	0.42	0.68	.484

the four subgroups within each of the two groups. The table also shows the *t*-values and probabilities.

Discussion. Inspection of Fig. 1 shows quite clearly that the relationship that was expected between the number of presentations of the list and the difference between the scores of incidental and intentional learning was obtained in this study. No difference between the scores of incidental and intentional learning was found after

only a few presentations of the list, but a difference was found after many presentations. The differences between the scores after 2, 6, and 8 presentations were not found to be statistically reliable. After 16 presentations of the list, however, the scores of intentional learning were found to be reliably higher than the scores of incidental learning (see Table I). The reliable difference between the two groups, which is found only after 16 presentations of the list, is a result of the increase in the scores of intentional learning in conjunction with the lack of an increase in the scores of incidental learning. None of the differences between the means of the subgroups in the incidental learning group was statistically reliable. On the other hand, statistically reliable increases in the scores of the intentional subgroups were found when the number of presentations of the list was increased (see Table II).

The previous study by Neimark and Saltzman found that the difference between the scores of incidental and intentional learning depends on the rate of presentation of the learning material: no difference between the scores was found at a fast (2-sec.) rate of presentation, but the intentional learning scores were significantly higher than the incidental learning scores at a slow rate (6 sec.).² The present study yielded similar results. The difference between the scores of incidental and intentional learning depends on the number of presentations of the learning material: no difference between the scores was found after few (2, 6, or 8) presentations of the learning material, but the scores of intentional learning were significantly higher than the scores of incidental learning after many (16) presentations. Summarizing the results of both studies, we may say that the difference between incidental and intentional learning depends on the total exposure-time of the stimulus-material, where total exposure-time can be manipulated either in terms of (a) the duration of each exposure or (b) the number of exposures.

² Neimark and Saltzman, *op. cit.*, 618-621.

APPARATUS

THE MEASUREMENT OF AUTOKINETIC MOVEMENT

By CECIL C. BRIDGES and M. E. BITTERMAN, University of Texas

In a recent review of the literature on autokinetic movement, Reinwald noted that the chief obstacle to systematic research on the problem has been the lack of a suitable technique of measurement.¹ Verbal estimations by *O* of the extent and direction of movement are obviously inadequate for any but the crudest purposes. Asking *O* to trace the course of movement as it occurs—a method first employed by Guilford and Dallenbach—represents a considerable advance over verbal description,² and the tracings are especially useful when a temporal frame of reference is provided by small 'jogs' of *O*'s pencil made in response to a regularly recurring signal.³ At best, however, the tracing method permits serious distortion of the pattern and the extent of movement. In one experiment by Reinwald, gross errors of size and shape appeared in the productions of *O*s who were shown a simple contour at a distance of 5 m. and asked to trace it on a piece of paper without looking at the pencil (in the same manner as the path of movement is traced in experiments on autokinesis). Reinwald also reported that trained *O*s often are surprised, after a period of autokinetic observation, at the marked discrepancies between the movement they have experienced and the tracings they have made.

The method reported here represents a quite different approach to the problem. The principle is simple: *O* is provided with a lever which permits him to vary the position of the autokinetic stimulus in the frontal-parallel plane; he is told that the point of light will be moved from its original position by a special mechanism and that he is to return it always to its starting point. The task is thus presented to *O* as a problem in tracking, and the course of autokinetic movement is indicated (inversely) by the readily recorded physical movement of the stimulus. It should be noted that even sophisticated *O*s, who know that the movement for which they

¹ F. L. Reinwald, *An Experimental Investigation of Conditions Affecting Autokinetic Sensations*, Doctoral dissertation, University of Texas, 1952, 1-199.

² J. P. Guilford and K. M. Dallenbach, A study of the autokinetic sensation, this JOURNAL, 40, 1928, 83-91.

³ Reinwald, *op. cit.*, 51-52.

are compensating is illusory rather than real, experience no difficulty in the use of this method.

The apparatus is diagrammed in Fig. 1. It consists essentially of a lever, pivoted in both horizontal and vertical planes, with a point-source of light (the autokinetic stimulus) on one end. The lever provides a 5:1 reduction in movement (*e.g.* *O* must move his end of the lever through a distance of 5 in. to produce a 1-in. movement of the stimulus), although

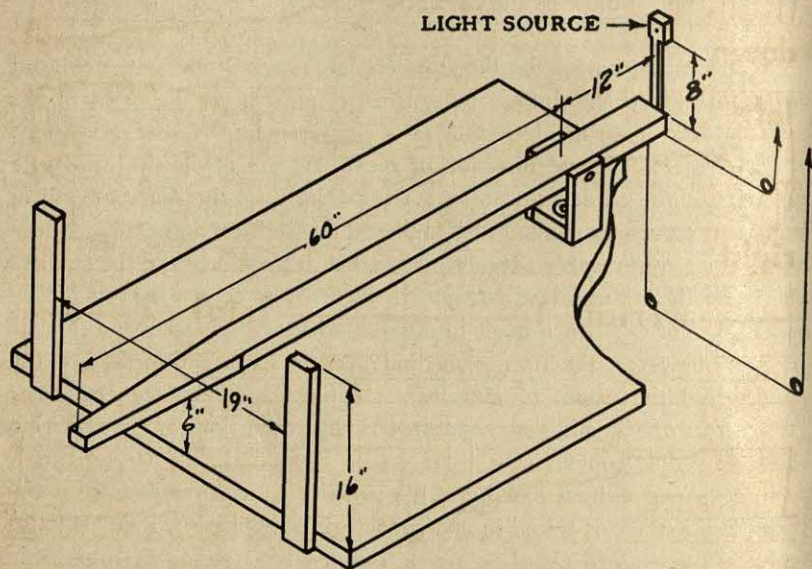


FIG. 1. DIAGRAM OF THE APPARATUS

any other suitable ratio may, of course, be used. The autokinetic stimulus is produced by a small neon lamp mounted behind a $\frac{3}{64}$ -in. hole in a small light-box. The box is set on a pillar which raises the point of light 8 in. above the top surface of the lever (to the eye-level of *O*). Stops are fixed at *O*'s end of the lever to limit its horizontal travel. To enhance (for the naïve *O*) the impression that the movement for which he must compensate is real, a small electric motor providing noise and vibration is attached to the lever just below the light, and a black fibreboard shield conceals the absence of a light-moving mechanism. (The motor and shield are not shown in Fig. 1.) Threads attached to the lever run to two pens of a continuous-feed kymograph which make simultaneous recordings of the horizontal and vertical components of movement. Sample tracings are shown in Fig. 2.

The suitability of the apparatus was examined in a simple study of the effect of angle of regard on the direction of autokinesis. A number of previous investigators have reported that autokinetic movement is in-

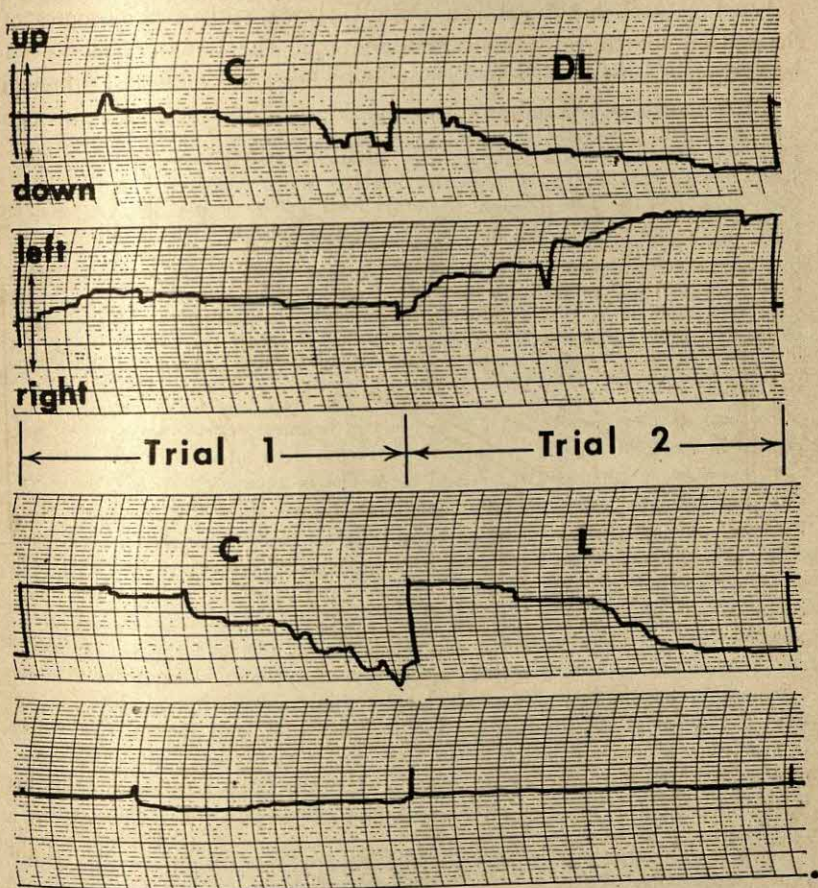


FIG. 2. SAMPLE RECORDS

The upper part of each record shows the vertical component and the lower half of each record the horizontal component of compensatory movement by *O* of the autokinetic stimulus. The upper record shows two consecutive 2-min. trials, the first with the head in the control (*C*) position and the second with the head tilted downward and to the left (*DL*). The lower record shows a control trial and a subsequent trial on which the head is tilted to the left (*L*). The autokinetic stimulus is recentered by *E* at the conclusion of each trial.

fluenced by the position of the eyes in the head, with the direction of the movement corresponding to the orientation of the eyes.⁴

⁴ H. A. Carr, The autokinetic sensation, *Psychol. Rev.*, 17, 1910, 42-75; H. F.

Observers. The *O*s were 16 undergraduate women drawn from a large University dormitory. They had no knowledge of the autokinetic phenomenon or of the purpose of the experiment.

Procedure. *O* was seated before the apparatus (in a chair of adjustable height) and her head was placed in a rest (made from a baseball catcher's mask from which the grill had been removed). The orientation of the head-rest could be varied on both the vertical and the horizontal axes. The following instructions were given:

This is an experiment on tracking and coördination with reduced visual cues. This mask will help you keep your head steady. When I turn off the overhead lights you will see a small red light here. Sometime after I turn off the overhead light, the red spot of light (here) will begin to move. Your task is to keep the spot of light in the same position it started from. Notice that if the light should move to the right, you could move it back to center by moving this stick to the right, like this (demonstrates). If it moved to the left, you could bring it back to center by moving the stick to the left. If it moved up or down, you could bring it back by moving the stick up or down. Remember, to bring the light back to center you must move your end of the stick in the same direction that the light has moved. Notice that you have to move the stick a rather large distance to make the light move a rather small distance. You may not be able to keep the light in the center at all times, but do the best you can.

For each *O* there were seven 2-min. trials in darkness separated by 2-min. rest-periods in the light. The first, fourth, and seventh determinations were controls—*O*'s head was not tilted or twisted, and the angles of regard (horizontal and vertical) were zero. On the remaining four trials (the second, third, fifth, and sixth), *O*'s head was tilted or twisted (or both) by adjustments of the head-rest, and opposing changes in the angles of regard were necessary if *O* was to fixate the autokinetic stimulus. There were four conditions of 'tilt': (1) vertical (up or down), (2) horizontal (right or left), (3) simultaneous vertical and horizontal tilts in directions the same as those employed individually in the first and second conditions, and (4) simultaneous vertical and horizontal tilts opposite in direction to those employed in the third condition. There were thus four possible combinations of these conditions—*U-R-UR-DL*, *U-L-UL-DR*, *D-R-DR-UL*, and *D-L-DL-UR*, where *U* means *up*, *D* means *down*, *R* means *right*, and *L* means *left*. Each combination of conditions was used for four *O*s, each of whom encountered the four conditions of the combination in an order given by one of the rows of a 4×4 latin square.

Results. The quantitative results of the experiment were derived from the measurement of graphic records like those illustrated in Fig. 2. Each record contains two tracings, the upper showing the vertical component of the compensatory movement of the light and the lower showing the horizontal component of that movement. At the start of each trial the autokinetic stimulus is centered and each pen is on the midline of its portion of the chart. Each fragment of record in Fig. 2 shows movement on two trials, the first a control (*C*) in which there is no tilt of the head

and the angles of regard are zero, and the second an experimental trial on which the head and eyes are tilted (in opposed directions). The upper record shows little movement on the control trial followed by marked movement downward and to the left when the head is tilted downward and to the left (*DL*). It should be noted that the record is labeled to show the relation between tilt of the head and compensatory (physical) movements of the light by *O*. Compensatory movements which correspond in direction to the tilt of the head support the hypothesis that autokinetic movement will correspond in direction to the tilt of the eyes in the head, e.g. with the head tilted *DL*, the eyes are *UR*, and when the compensatory movement is *DL*, the autokinetic movement is *UR*. The upper record thus provides support for the hypothesis under scrutiny. The lower record, however, does not: On both trials (upper tracing) there is pronounced compensatory movement in the downward direction which is understandable since the change in conditions from the first trial to the second was not expected to effect the vertical component of movement; but the tilt of the head to the left had no great effect either on the horizontal component of movement (lower tracing).

Each tracing was measured with a compensating polar planimeter which provided an index of the direction and extent of its deviation from the starting point.⁵ For each *O* the measures derived from the three control tracings were averaged to provide vertical and horizontal baselines with which to compare the movement occurring under experimental conditions. The measures for each experimental trial were subtracted from the basal measures; the magnitude of the resulting difference indicated the effectiveness of the particular condition of tilt used on that trial, and its sign (*plus* or *minus*) indicated whether the effect was in the direction predicted by the hypothesis or in the opposite direction. The difference-scores of each *O* were summed algebraically to provide a single measure of the effectiveness of the experimental conditions for that *O*. The mean of these measures was significantly positive beyond the 1% level of confidence ($t = 3.63$ with 15 *df*). The results may therefore be said to confirm the hypothesis tested.

In general, both logical considerations and the results of preliminary tests suggest that the method here described is eminently suitable for systematic, quantitative work on the problem of autokinetic movement.

⁵ Manufactured by Keuffel & Esser.

APPARATUS NOTE

THE USE AND CALIBRATION OF THE 'MAXWELLIAN VIEW' IN VISUAL INSTRUMENTATION

The radiant energy reflected by or emitted from a surface diverges in all directions.¹ In order to be effective visually, this energy must pass through the entrance pupil of the eye after which, assuming adequate dioptrics, it will be focused on the retina. The size of the light pencil emanating from a given object point which reaches the corresponding image point on the retina is proportional to the pupil diameter. The larger the pupil diameter (or the light pencil) the more energy from the object point will be able to reach its corresponding image point and the 'brighter' the point will appear. The entrance pupil of the eye can, however, intercept only a small fraction of the energy from a diffuse surface, and consequently most of the emitted energy does not enter the eye and is therefore not effective visually (Fig. 1A).

In order to present an extended test-field of high luminance, the 'Maxwellian view'

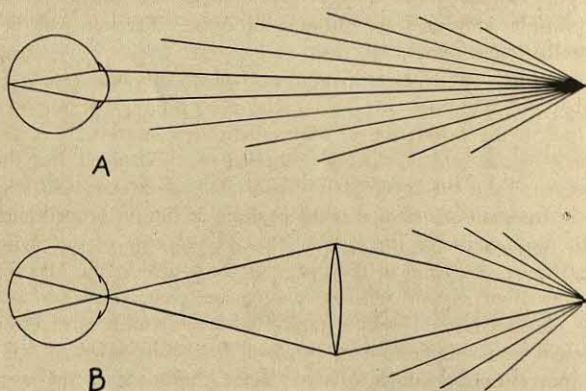


FIG. 1. DIAGRAM ILLUSTRATING THE EFFICIENCY OF THE 'MAXWELLIAN VIEW' A, unaided eye, only the light intercepted by the pupil reaches the retina; B, Maxwellian view, all the light intercepted by the lens enters the eye.

is frequently used in visual instrumentation.² The principle of the Maxwellian view is to focus the image of a light source in the plane of S's pupil (see Fig. 1B). All the light intercepted by the lens enters the eye as contrasted with Fig. 1A where

¹ The author wishes to express his appreciation to Dr. C. S. Bridgman for his effective aid in the preparation of this note.

² Named in honor of J. Clerk Maxwell who utilized this principle in the construction of his color-mixing apparatus. (See Maxwell, On the theory of compound colors and the relation of the colors of the spectrum, *Phil. Trans.*, 150, 1860, 57.)

only the light intercepted by the pupil enters the eye. Subjectively, *S* sees the lens (called the field lens) filled with light. With the aid of additional lenses the light can be rendered parallel hence, neglecting absorption in the lens and surface reflections, it can transverse long optical paths without loss. Variations in temporal characteristics, direction, and energy may be produced by modification of this beam by the appropriate shutters, lenses, reflecting surfaces, filters, etc.

An advantage of the Maxwellian view, in addition to its efficiency in transmitting energy from the source to *S*'s eye and the possibilities it offers for manipulation and control of the light, is the elimination of fluctuations in natural pupil size as an experimental variable provided the image of the source in the plane of the pupil is smaller than *S*'s natural pupil. In addition, since light from each point of the source passes through all zones of the lens, irregularities in the source will not affect the uniform appearance of the field lens.³ It is not advisable, however, for the source to exhibit gross differences in luminances such as would be obtained from a coiled tungsten filament unless the image of the filament is so small as not to be imaged on the edge of the pupil. If the latter situation exists, movements of the eye can result in variations in the total flux reaching the retina as the pupil edge moves across areas of non-uniform flux-density. A disadvantage of the Maxwellian view is the necessity for placing the eye in such a position that the image of the source is centered in the entrance pupil, although this arrangement does permit the presentation of peripheral stimuli without decrease in the effective area of the entrance pupil.⁴ Another disadvantage may arise in experiments on visual acuity where the possible additional complication of Abbe diffraction may be introduced by the Maxwellian view optics.⁵

Since most visual photometers (e.g. the MacBeth illuminometer) are designed for the calibration of diffuse surfaces, special precautions must be observed if they are to be used in the calibration of Maxwellian view systems. It is necessary, that the appearance of the field lens being calibrated is not changed, that the exit pupil of the photometer and the image of the source in the Maxwellian system be concentric and in the same plane so that the position of the eye is the same for calibration as in the normal use of the system. The aperture-stop of the photometer must also be equal to or smaller than the image of the source in the Maxwellian system that the Maxwellian system and the photometer comparison surface be viewed through the same diameter entrance pupil. This requirement is necessary to satisfy the requirement that specification of luminance be independent of viewing conditions. It is not always possible to satisfy these conditions as the eyepiece of the apparatus may be so constructed as to prevent the proper positioning of the photometer or the aperture-stop of the photometer may be too large. In some cases it may be possible temporarily to alter the design of the apparatus or the photometer so as to satisfy these requirements (e.g. by the use of a supplementary lens which

³ If the image of the source in the plane of the pupil is too small, entoptic phenomena will destroy the even appearance of the field lens. See A. A. Michelson, On the effect of small particles in the vitreous humor, *J. Opt. Soc. Amer.*, 9, 1924, 197.

⁴ See the apparatus used by R. T. Brooke, Variations of critical fusion frequency at various retinal locations, *J. Opt. Soc. Amer.*, 41, 1951, 1010.

⁵ E.g. the apparatus designed by S. Shlaer, The relations between visual acuity and illumination, *J. Gen. Physiol.*, 21, 1937, 165.

forms a conjugate image of the source with unit magnification at some distance away from the eyepiece of the apparatus). Alternative procedures for the calibration of the Maxwellian view which permit the photometer to be used on diffuse surfaces are indicated below.

A convenient technique for calibration of the Maxwellian view is to cover one half of the field lens with a mirror positioned so as to reflect the image of a diffuse surface in the direction of S's eye. S then sees a bipartite field, one half of which is supplied by the Maxwellian view field lens and the other half by the image of the diffuse surface seen in the mirror. The two halves of the field lens are matched by adjusting the luminance of either component after which a calibration value is obtained in the usual manner from the diffuse surface. As in the calibration method discussed above, the effective entrance pupil must be the same size for the Maxwellian and the diffuse fields.⁶

In the binocular method S observes the field of the Maxwellian view with one eye and a diffuse surface of the same size with the other eye. By adjusting convergence, the two fields may be made to appear in juxtaposition. After being matched for equality, the luminance is determined with a conventional photometer from the diffuse surface. It is necessary, of course, that the entrance pupils of the two eyes be equal. The facility with which this binocular match can be made is improved by providing for a common boundary between the two fields. This can be accomplished by presenting both fields in the form of a semi-circle or rectangle. The variability of binocular matches is, however, greater than for monocular comparison.

A third technique makes use of a photo-electric circuit.⁷ The sensitive surface of the photocell is placed in the plane normally occupied by S's pupil and the voltage generated by the image focused there is recorded. A diffuse surface equal in size to the field lens is illuminated and the light reflected into the photocell located at a distance equal to that between the eye and the field lens. The area of the photocell exposed to this light must be equated with the area of the image of the source previously focused on the photocell. The illumination on the diffuse surface is varied until the flux density of the light reflected on the photocell surface, as indicated by the voltage generated there is equal to the flux density previously produced there by the Maxwellian view system. Under this condition the effective energy entering the S's pupil in the Maxwellian system is equal to that which would enter the pupil from the diffuse source. The luminance of the diffuse source can then be determined in the conventional manner.

Summary. The principle of Maxwellian view is advantageously used in visual experimentation when it is desirable to obtain extended visual stimuli of high luminance. Precautions in the use of this system as well as some advantages and disadvantages are discussed. Calibration of the system involves special problems when using conventional photometers. Alternative methods for calibration which make use of monocular comparison, binocular comparison, and a photoelectric circuit are described.

University of Wisconsin

HERSCHEL LEIBOWITZ

⁶ The author is indebted to Dr. Yun Hsia for the details of this technique.

⁷ Described by K. J. W. Craik, The effect of adaptation on differential brightness discrimination, *J. Physiol.*, 92, 1938, 406.

NOTES AND DISCUSSIONS

AN OPTICAL ILLUSION OF SPONTANEOUS FLUCTUATIONS IN APPARENT SIZE

The illusion reported here was first observed by the author in an experimental apparatus. His head was held steady by a chin rest and forehead pads and a white field was observed monocularly through an aperture. The aperture, a round hole ($\frac{1}{4}$ in. in diameter) in a piece of sheet metal, was placed 3 in. from his eye and the observational field, 27 in. beyond the aperture. A fixation-point in the observational field was being regarded steadily through the center of the aperture when the illusion was first noted. The illusion consisted of small apparent expansions and contractions of the aperture, which seemed randomly distributed with respect to time. At least 10 Os have seen the illusion since and only one has been found who has not seen the illusion clearly.

This illusion may be explained in terms of fluctuations in pupillary diameter. It is well known that even under conditions of steady stimulation there are small fluctuations in pupillary diameter.¹ These arise from the antagonistic action of the sphincter pupillae and the dilator pupillae, which control the diameter of the pupil. Furthermore, when the eye regards a field through an aperture, the size of the field which can be seen is determined partly by the size of the pupil. This is illustrated in Fig. 1. If the minor effect of refraction at the cornea is disregarded, then the limiting point on the right side of f , the field which can be seen, is determined by a straight line drawn from the left side of the pupil, p , past the right side of the aperture, a . The reason the limiting point is determined in this way is that light from any point farther to the right, misses the pupil entirely. Of course the left limit of the field seen is determined in a corresponding manner. It is obvious from Fig. 1 that if a , x , and y are held constant, any change in pupillary diameter will change the size of the field seen. Thus it may be expected that under the conditions described there should be spontaneous fluctuations in the size of the field seen, and hence in the visual angle subtended by that field. It may be noted from Fig. 1 that f and a have identical projections on the retina, so the visual angle subtended by the aperture is identical with the visual

¹ F. H. Adler, *Physiology of the Eye; Clinical Application*, 1950, 148.

angle subtended by the field seen. Thus there are also spontaneous fluctuations in the visual angle subtended by the aperture. Since the observer is entirely unaware of the fluctuations in pupillary diameter, which cause the fluctuations in visual angle of the aperture, this is theoretically a perfect *reduction* situation. Hence apparent size of the aperture should follow

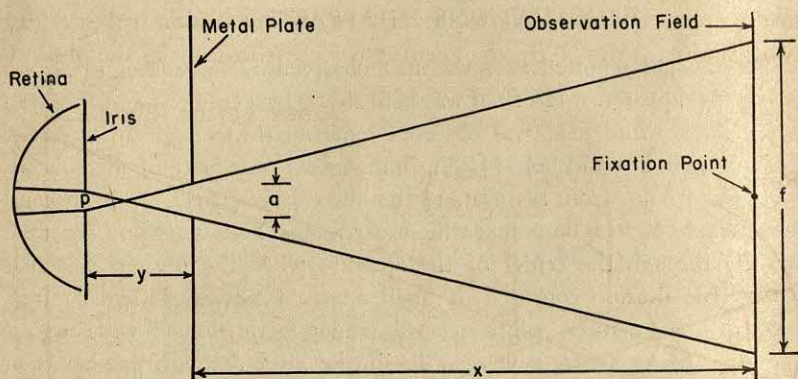


FIG. 1. VIEW FROM ABOVE OF OPTICAL SYSTEM IN WHICH THE ILLUSION WAS FIRST OBSERVED

perfectly the change in visual angle subtended by the aperture, and there should be spontaneous fluctuations in apparent size of the aperture.

There are several points in this theoretical explanation of the illusion which may require clarification. The visual angle subtended by an object is ordinarily determined by the physical size of the object and its distance from the nodal point of the eye. In observing this illusion both the physical size of the aperture and its distance from the eye are fixed, so if the visual angle of the aperture were determined in the usual way this angle could not vary. The usual method of measuring visual angles, however, applies primarily to those cases where light from each point on the object fills the pupil, and where the eye is accommodated for the object. In the present case all light through the aperture comes from the observational field, which is considerably more distant than the aperture. Moreover, the light which passes from each point on the limit of the field seen, past the edge of the aperture, is a marginal ray; in other words, these rays do not pass through the nodal point of the eye. Since the eye is accommodated for the observational field, the point where each of these marginal rays stimulates the retina is determined by a straight line drawn from the origin of the ray, at the limit of the field seen, through the nodal point of the

eye to the retina. This is illustrated by Fig. 2. Thus the visual angles subtended on the retina by both the field seen, and the aperture, are given by angle b in Fig. 2. Since these angles are identical, any change in size of the field seen will cause identical variation in the visual angles subtended by both the field seen and the aperture.

One may wonder why the changes in apparent size should be referred to the aperture, rather than to the field itself. There seem to be at least two reasons for this. First, the aperture is seen as in front of the observa-

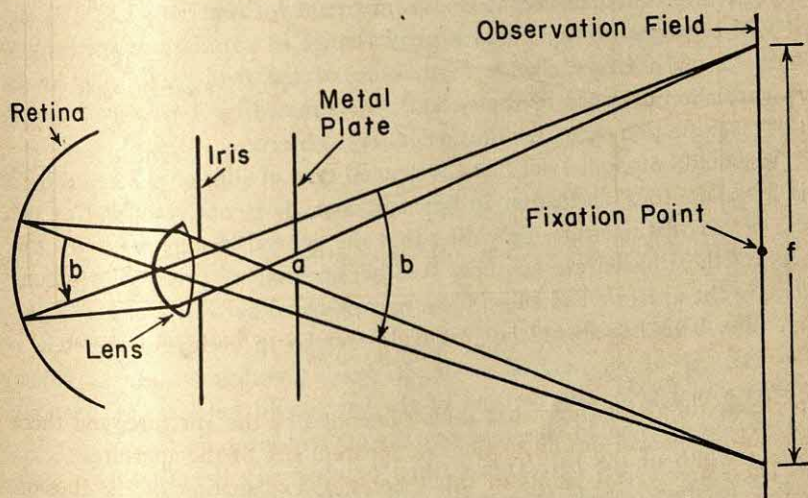


FIG. 2. DRAWING SHOWING METHOD FOR DETERMINING THE VISUAL ANGLE, b , SUBTENDED BY BOTH THE FIELD SEEN, f , AND THE APERTURE, a

tional field and as limiting the amount of the field seen. Under these conditions the only way O knows for the visual angles of aperture and field to vary concomitantly is by way of changes in size of the aperture. Secondly, change in visual angle of the field is accompanied by seeing a greater or lesser amount of the field. There are no changes in visual angle of a given object or distance in the field, and hence no changes in apparent size of the field itself.

One may also wonder why the illusion reported here is not commonly observed. The reason is that good conditions for observing the illusion are seldom encountered. Both the head and the instrument should be held so steadily that relative movement between head and instrument does not make the changes in apparent size of the aperture difficult to observe. The aperture should subtend a visual angle so small that the changes in

apparent size can be seen easily, and it should be far enough from the eye that it does not act merely as an artificial pupil. *O*'s eye must be accommodated for the observational field rather than for the aperture itself. If the eye is accommodated for the aperture, then an increase in pupillary diameter increases the amount of light passing the edge of the aperture which enters the eye, but all of this additional light comes to a focus at the same point on the retina. Thus there is an increased brightness at the edge of the aperture but no change in visual angle subtended by the aperture. The distance, x , is also important for observing the illusion. It will be noted in Fig. 1 that a small change in pupillary diameter produces a much larger change in the size of the field seen. The linear magnification obtained by the system is determined by x/y , and of course this magnification makes the illusion easier to observe.

It should be pointed out that the general type of illusion described here is a spontaneous fluctuation in apparent size. It is not restricted to the use of a visual aperture, providing that the general requirements for observing the illusion are satisfied. Another method of producing it is to remove the aperture and substitute a narrow metal bar in the plane of the aperture. This bar should be just wide enough to occlude completely a narrow strip of the field of vision. It should be aligned along a radius of the visual field, and the end of the bar should be near the fixation-point. The illusion in this case consists of very clear fluctuations in apparent width of the bar. This method of observing the illusion is even better, for two reasons, than the use of an aperture: (1) the bar can be brought closer to the fovea than the edges of the aperture; and (2) the visual angle of the bar can be made much smaller than the visual angle of the aperture, but for a given change in pupillary diameter the change in visual angle of the image is approximately the same for both bar and aperture. This means that the changes in visual angle of the image, due to changes in pupillary diameter, are more striking in the case of the bar. It should be noted that the relation between pupillary diameter and apparent size is opposite with the bar from what it is with the aperture. An increase in pupillary diameter increases the apparent size of the aperture, but decreases the apparent size of the bar. The reason for the relation between pupillary diameter and apparent size of the bar is shown in Fig. 3, in which f is the part of the observation field which cannot be seen.

The above considerations seemed to give a complete theoretical account of the illusion, and the interpretation was easy to test. First, observations were made to see if changes in pupillary diameter have the effect on apparent size of the aperture which is required by the interpretation. To do this with maxi-

mal clarity, the consensual light reflex was used. This is the reflex whereby a stimulus to one eye produces pupillary contraction in the other eye, and contrawise, its removal produces dilation in the other eye. Use of this reflex has the advantage that one eye can be stimulated, while the other eye is observing continuously to watch the rapid changes in pupillary diameter which occur. It also has the advantage that dilation, as well as contraction, can be observed. The results of these observations were quite striking. When the light was turned on the apparent size of the aperture quickly decreased, and when the light was turned off the apparent size of the aperture increased. Apparent expansion of the aperture occurred more slowly than apparent contraction, corresponding to the slower speed of pupillary dilation.² When the pupil undergoes a large change in diame-

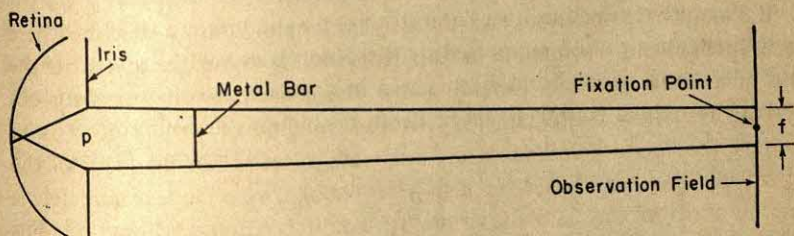


FIG. 3. VIEW FROM ABOVE OF OPTICAL SYSTEM FOR OBSERVING THE ILLUSION WITH A VERTICAL METAL BAR

ter, it does so by successive steps.³ Pupillary contraction is so rapid that it proved difficult for the *Os* to see these steps during contraction. During dilation, however, from three to five steps of increase in apparent size of the aperture were seen clearly. All of these observations were made by a number of *Os* with complete consistency of results.

These observations made it clear that changes in pupillary diameter have the effect on apparent size of the aperture which is required by the foregoing interpretation. Thus any change in pupillary diameter will give the illusion of a change in apparent size of the aperture. The question remained whether the spontaneous fluctuations in apparent size of the aperture are in fact caused entirely by spontaneous fluctuations in pupillary diameter. This question was answered by making observations with and without an artificial pupil. The artificial pupil was 2 mm. in diameter, and hence was smaller than the natural pupil under the conditions of observation. Thus with the artificial pupil any effects of spontaneous

² W. S. Duke-Elder, *Text-Book of Ophthalmology*, 1, 1942, 554-555.

³ *Ibid*, 555.

fluctuations in pupillary diameter were removed. A number of *Os* were used who saw the spontaneous fluctuations in apparent size of the aperture quite clearly without the artificial pupil. All of them reported that the spontaneous fluctuations disappeared entirely when the artificial pupil was used. Of course an artificial pupil not only fixes the size of the pupil; it also makes the field seen smaller and darker. The two latter conditions were produced independently by reducing the size of the aperture and reducing the illumination on the observational field, but under these conditions the spontaneous fluctuations in apparent size of the aperture were not reduced. Thus the artificial pupil must have removed all spontaneous fluctuations in apparent size of the aperture because it removed the effects of spontaneous fluctuations in pupillary diameter.

All of these observations to test the theoretical interpretation of the illusion were repeated using a bar instead of an aperture, and the theoretical interpretation was also supported by the observations with the bar. Hence the illusion of spontaneous fluctuations in apparent size, as reported here, may be attributed entirely to spontaneous fluctuations in pupillary diameter.

Aero Medical Laboratory

KENNETH T. BROWN

Wright Air Development Center, Ohio

AUDITORY FATIGUE BEYOND AND WITHIN THE COMPASS OF THE VOICE

When auditory fatigue is produced with a stimulus-frequency that lies considerably beyond the compass of the voice, it appears that the maximal effect is noticeably *shifted* from the stimulus-frequency upwards. Some time ago Dr. Onchi sent me a brief report of experiments he had made in which he exposed persons of normal hearing to continuous stimulation by a pure tone of 1 dyne/cm² sound-pressure for 5 min. and quickly thereafter measured their auditory fatigue in terms of a rise in threshold.¹ He asked for a theoretical explanation of the shift observed. The explanation given below is mine, but the experimental data are his.

To simplify the theoretical argument, I use only two stimulation-frequencies. They are illustrated by the *broken* curves shown in Figs. 1 and 2. These figures show in decibels the amount of intensification necessary to overcome the fatigue. Seemingly auditory fatigue beyond the compass of voice-frequencies is not simply electro-chemical or neural fatigue but a compound of that and a mechanical hindrance which I think is

¹ Dr. Yutaka Onchi, 1936 Takadacho, Kohoku-ku, Yokohoma, Japan.

produced by bruising tissues and thus for 20 min. (or more) reducing their flexibility.

I have for a long time nursed the conviction that the really 'regular' functioning of the human cochlea is confined within frequencies barely surpassing the compass of tone production by human larynxes, which means from about 90 to about 1000~. From the viewpoint of a biologist it is plausible that there should be a relation between a certain limitation of functioning of the cochlea and the limited functioning of the motor organ doing the signaling to reach the ear of other individuals of the species. Guided by this belief in an evolutionary relation, I chose for discussion the frequency 600~, lying within the human voice, and another frequency, 4000~, lying quite outside the compass of the human voice.

Fig. 1 illustrates the experiment with the frequency of 600~. The fatigue is shown in the drawing as if it were something positive, amounting to 10 db. for the testing frequency 600~ and falling off at the sides of 600 in the manner of a probability curve, extending over no more than ± 150 ~.

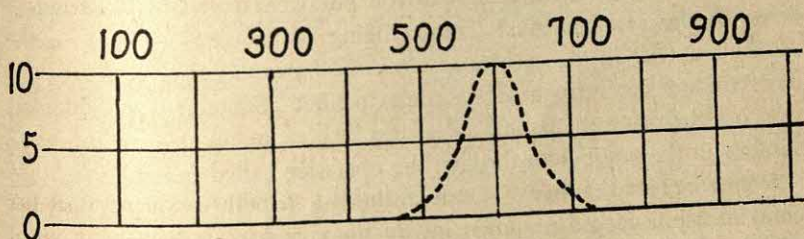


FIG. 1. AUDITORY FATIGUE WITH THE STIMULATING FREQUENCY WITHIN THE COMPASS OF THE HUMAN VOICE

I now ask the questions: Where can that fatigue be located physiologically? Of what nature is it? I find it most plausible to compare it with the fatigue occurring in a cone of the retina. Nothing is easier than to demonstrate that a place on the retina, after lengthy exposure to green light, becomes less sensitive to this and neighboring wave-frequencies, but still responds with undiminished intensity to blue light and to yellow light. Led by this analogy, I regard it as most probable that *this* auditory fatigue (*i.e.* threshold-rise but not other loudness) is of an electro-chemical nature and located in the haircells of the phragma and possibly also in the affected fibers of the auditory nerve. This then does not immediately propose any additional problem.

The experimentally measured auditory fatigue caused by frequencies

definitely outside of the range of the human voice gives us in Fig. 2 a picture of a problem. If the measured fatigue for the frequency 600~ would again be found for the frequency 4000~, there would be no *shifting* of the maximum from 4000~, where the stimulation occurred, to a point approaching 4500~. (The change from 10 db. to 40 db. is a minor matter which might be left out of the discussion.) A similar shift to the right was observed by the experimenter in other cases lying above a stimulation frequency of 1000~.

In explaining the difference between the results shown in the broken curves of Figs. 1 and 2, I assume that the cochlea is a hydraulic system activated by a kind of hydraulic ram which is represented by the liquid in the vestibulum (not in the cochlea). The vestibular liquid, in consequence of its inertia, swings toward and away from the beginning of the cochlea, where the round window is located, whenever the skull is shaken. If the round window did not exist, no such swinging of the vestibular liquid could occur no matter how strongly the skull were shaken. If the liquid is to swing and find an outlet (or inlet) at the round window, a certain length of the phragma must be *jerked* up, or down as the case may be.² The swinging of the oval window in the vestibulum (not in the cochlea as often misdrawn) is also necessary. The oval window not infrequently becomes ossified. Contemporary otological surgeons then cure the patient's deafness by creating an artificial substitute window in the vestibulum through the operation called fenestration.

I proceed on the basis of the following hypotheses: (1) that the phragma while jerked up-down *within* the compass of the human voice is not easily bruised; after 5 min. its flexibility remains normal; but with frequencies beyond the human voice it is apt to become bruised and swollen and, like a region of inflamed human skin, to lose to a degree its flexibility and (2) small up-down jerks of a part of the phragma may leave the sensitive cells within it uncharged hence intense jerks are required to charge the cells.

We first apply Hypothesis 1. While the hydraulic ram of the vestibular liquid between the oval window and the cochlea entrance swings incessantly for 5 min. in so high a frequency as 4000~, a piece of the phragma becomes bruised. What piece and where? The case is similar to the damage done near the 'eye' of a hurricane. The cochlea entrance is the 'eye.' The nearer the 'eye,' the greater the damage. In a hurricane the damage depends also on the wind velocity near the 'eye.' This is analogous to the high

² Cf. M. F. Meyer, *How We Hear*, 1948, 16.

frequency (now 4000~) of jerking the phragma. We answer the question 'where' by saying the bruised piece is near the cochlea entrance, near the round window.

We now consider Hypothesis 2. When the 'fatigue-test' is used, we find that the sensitive cells near the round window are unchargeable because of the phragma's reduced flexibility. When, however, the amplitude of the testing tone is increased until it reaches a *region* of the phragma which has lost little of its flexibility, its sensitive cells are charged.³ Thus 10 db. grow to 40; but what about the *shift*, which is our real problem? The magnitude of bruising the tissues and *reducing their flexibility as depending purely on the frequency* of the damaging 5-min. tone is illustrated in Fig. 2 by the steadily rising curve beyond the voice-compass. The figure tells that very little damage is done by frequencies below

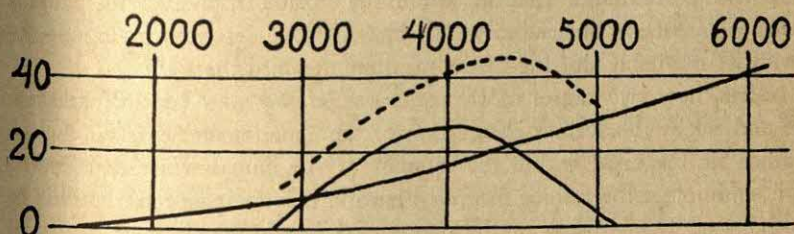


FIG. 2. FLEXIBILITY-REDUCTION OF THE PHRAGMA BY FREQUENCIES ABOVE 1000~. The long solid curve is the hypothetical flexibility-reduction by various frequencies varying from 1000 to 6000~. The solid symmetrical curve is the hypothetical effect of electro-chemical fatigue. The broken curve is the effect, experimentally determined, of test-stimuli beyond the compass of the human voice.

1000~. More damage is done by the frequencies around 2000~, still more damage by 3000~, by 4000~, and so on. We recall that wind velocities (without reference to the 'eye') vary in different hurricanes and the damage varies.

Now the experimenter tests the auditory fatigue for 4000~ (or rather, what naïvely might be called fatigue) by applying sufficiently high-powered stimuli from 3000 to 5000~. He expects a symmetrical rise of 'fatigue' to a maximum at 4000~ and a fall beyond that point. But there is no symmetry, and the maximum is not at 4000~ but *shifted* toward the right!

After all that we have said about the voice compass, the shift toward the right in Fig. 2 appears natural. Any bruised region is the less flexible the greater the frequency which tries to flex it in the test. Try it on a

³ Cf. Meyer, *op. cit.*, Fig. 4, p. 16.

piece of rubber plate, say, a piece of tire tube. You can flex it easily once or twice per second, but with increasing difficulty five or ten times per second. So, when the tester flexes the bruised phragma region only 3000 times per second, he finds a flexible-enough region *not too far* from the entrance to the cochlea and obtains a 'hearing' response. But (as the rising curve indicates) when he tries 4000 or 5000 flexes per-second, he has to *increase the amplitude of the test tone* to reach a phragma region where through better preserved flexibility (farther from the 'eye') displacement-'jerks' can be produced sufficiently strong to charge the sensitive cells (the electro-chemical fatigue not being considered at this moment of the discussion).

The total effect then is a compound of the symmetrical curve of the true (*i.e.* electro-chemical) auditory fatigue and the unsymmetrical curve of the mechanical hindrance resulting from the inflammation of the phragma with its reduction of flexibility. The broken curve describes the compound with its inevitable shifting of the maximum toward the right.

That these two causes of the *apparent fatigue* must be distinguished, is verified by the further observation of the experimenter that the fatigue which he measured within the compass of the human voice disappeared in 5 min.; but the fatigue measured beyond that compass took 20 min. to disappear. The bruising of the phragma, I should say, is not relieved as quickly as the electro-chemical loss of sensitivity.

The above discussion might also be applied to a critique of a historical matter. During the last fifty years several authors reported that they had used incessant stimulation with high frequencies (2000~ or more) for hours and even days on animals, and that *post mortem* examination yielded the following tissue destruction in the phragma: (1) the greatest destruction was found near the cochlea entrance (the hurricane's eye, in my analogy) rather than farther away; (2) the greatest destruction was found for stimulation with higher rather than lower tone frequencies (greater wind velocities of the hurricane).

From combining these two premises they drew the conclusion that *Nature had tuned the phragma* like a series of piano strings, and that the destruction was due to *excessive resonance*. (For undisclosed reasons the hydraulic conception of cochlear mechanics was repugnant to them and, though offered to them, was left completely out of consideration.)

It does not seem to me that their conclusion was either compelling or anatomically plausible. They never tried the crucial experiment of using very low frequencies (say, between 100 and 150~) for experimentation

with the expectation of finding tissue damage exclusively near the apex where they *imagined tuning* for the lowest frequencies.

Miami, Florida

MAX F. MEYER

CONSTANT ERRORS IN THE MEASUREMENT OF FIGURAL AFTER-EFFECTS

Within several months last year, two papers on bilateral constant error in size judgment were published. Brown measured differences in the apparent length of lines of physically equal size, when exposed in two different halves of the visual field.¹ He called such a difference a half-meridional difference (*HMD*). McPherson and Renfrew found that when physically equal-sized small metal circles are felt simultaneously by the two hands of a blindfolded *S*, the object held in the preferred hand is usually judged smaller.² Following a suggestion of Brown's, one could call such a difference a bilateral kinesthetic difference (*BKD*).³

The results obtained by McPherson and Renfrew are clearly relevant to studies of the kinesthetic figural after-effect, in the same way as Brown's are important for the methodology of studies of the visual figural after-effect.⁴ Just as the figural after-effect and the *HMD* in vision are methodologically quite easy to distinguish, as Brown has pointed out, so the kinesthetic figural after-effect and the *BKD* are operationally quite distinguishable. *S* need only make a judgment of subjective equality of size before the inspection-pattern is presented. Any divergence between the *PSE* thus obtained and true physical equality will be a measure of the *BKD*. Then the inspection-pattern is presented, and *S* is again asked to find the width on the right-hand bar which feels just as wide as the one in his left hand. The difference between this latter match and the pre-inspection match will be a measure of the size of the figural after-effect. Thus, in using Köhler and Dinnerstein's technique for measurement of the kinesthetic figural after-effect,⁵ it is not unlikely, as Köhler and Dinnerstein point out,⁶ that *BKDs* may affect the measurement (even

¹ K. T. Brown, Factors affecting differences in apparent size between opposite halves of a visual meridian, *J. Opt. Soc. Amer.*, 43, 1953, 464-472.

² A. McPherson and S. Renfrew, Asymmetry of perception of size between the right and left hands in normal subjects, *Quart. J. Exper. Psychol.*, 5, 1953, 66-74.

³ Brown, personal communication.

⁴ Brown, Methodology for studying figural after-effects and practice effects in the Müller-Lyer illusion, this JOURNAL, 66, 1953, 629-634.

⁵ Wolfgang Köhler and D. Dinnerstein, Figural after-effects in kinesthesia, in *Miscellanea Psychologica A. Michotte*, Louvain, 1947, 196-220.

⁶ *Ibid.*, 202-203.

though, in some preliminary work, they did not find any very strong *BKD*s). If Köhler and Dinnerstein's inspection-pattern is used on right-handed *Ss* in such a way as to make the right kinesthetic extent subjectively narrower than the left one, the total size of the discrepancy from physical equality will presumably, on the average, be larger, since the figural after-effect will summate with the *BKD* and both go in the same direction.

Data on kinesthetic figural after-effects including control judgments have been obtained in connection with another study.⁷ In addition to measuring the figural after-effect with the suggested control, these data also provide a check of McPherson and Renfrew's findings, which seems pertinent since, in their discussion, they write:⁸

The comparison of our results with those of previous investigators is not easy, for the methods employed are not strictly comparable. Bloch's method of testing perception of size by holding the pages of a book between thumb and forefinger is indeed the only one in which a comparison is reasonable. We have not been able to trace his original paper, and Sherrington's statement that Bloch found "... distance is underestimated with the left hand in right-handed people, but by the right hand in left-handed people," is so much at variance with our results that the suspicion of a misquotation is raised in our minds.⁹

In the course of the preliminary experiments with figural after-effects mentioned above, it was noticed that there seemed to be some constant errors in pre-inspection judgments, and consequently control measurements preceding the inspection-period were inserted as a regular part of the methodology for studying kinesthetic figural after-effects.¹⁰ The procedure was to request four average-error equality judgments of the test-pattern, as the first step in measuring the size of the figural after-effect. An estimate of handedness was available on the *Ss* studied, since in other parts of the procedure they were required to write with a pencil. Since the *BKD* is related to handedness, according to McPherson and Renfrew, one should be able to predict the direction of the constant errors occurring in the control measurements made on the kinesthetic apparatus.

⁷ Michael Wertheimer and Nancy Wertheimer, A metabolic interpretation of individual differences in figural after-effects, *Psychol. Rev.*, 61, 1954, 279-280; Michael Wertheimer, Figural after-effects in schizophrenic and normal subjects, *Amer. Psychol.*, 9, 1954, 492 (abstract); The differential satiability of schizophrenic and normal subjects: A test of a deduction from the theory of figural after-effects, *J. Gen. Psychol.* (in press); Figural after-effect as a measure of metabolic efficiency, *Quart. J. Exper. Psychol.* (in press).

⁸ McPherson and Renfrew, 73.

⁹ Bloch, quoted by McPherson and Renfrew from Sherrington in Schafer's *Text-book of Physiology*, 2, 1900, 1018; original not seen.

¹⁰ In this connection see G. S. Klein and David Krech, Cortical conductivity in the brain-injured, *J. Personal.*, 21, 1952, 118-148. These authors also used a set of pre-inspection control measures in determining the size of kinesthetic figural after-effects; Brown, too, as have other recent investigators, has worked with such controls.

The apparatus was copied from that of Köhler and Dinnerstein. *S*, blindfolded, was placed between two sawhorses, 24 in. apart, on which were placed, at a comfortable height, four wooden bars; there were two such bars end to end on each horse. The left horse held a near bar $\frac{1}{2}$ in. wide (inspection-pattern) and a far one $1\frac{1}{2}$ in. wide (test-pattern). The right horse held a near bar $2\frac{1}{2}$ in. wide (inspection-pattern) and a far one, which was 1 in. wide at the near end and increased in width by a total of 30 steps, each representing a discrete increase of $\frac{1}{16}$ in., to $2\frac{7}{8}$ in. wide at the far end (test-pattern). All bars were 60 in. long and 2 in. high.

Standing between the far (test) bars, *S* was requested to "find the place on the right-hand bar which feels just as wide as the bar in your left hand seems to feel," using just the top edges of the bars and the tips of his fingers (thumb on one side, fingers on the other). After such a judgment *S* was moved back several feet, turned around, and then brought back approximately to the same position at which he had stood before, facing the same way. Then he was asked to make a second match. After a total of four such judgments, *S* was led back between the inspection-bars, and requested to rub back and forth on them, with "long, sweeping movements," for 20 sec. Then he was again led between the test-bars, and another judgment of equality was requested. There immediately followed another 20-sec. inspection-period, another test, and so on, until a total of four post-inspection judgments had been obtained. Throughout, *S* was asked to make his post-inspection judgments as rapidly as possible, and no *S* took more than 15 sec. nor less than 5 sec. in coming to a decision. Available for each *S*, then, were a mean of four control judgments and a mean of four experimental judgments, as well as the true physical equality point on the right-hand test-bar.

A total of 60 *Ss* was used: 33 men—29 right-handed and 4 left-handed; and 27 women—22 right-handed and 5 left-handed. About half were University students and their wives, and half hospital staff members.

The steps on the right-hand test-bar were numbered from 1 to 30, with step 9 physically the same width as the left-hand test-bar.

For the right-handed *Ss*, the average control choice was 11.18, with a standard error of 0.29. The difference between this value and physical equality (9.00) is significant, by *t*-test, at far better than the 0.001-level, clearly corroborating McPherson and Renfrew's finding. Curiously, the right-handed women tended to underestimate the right-hand width more than did the right-handed men. The control values were 12.49 and 10.34 for the women and men, respectively, a difference significant at better than the 0.01-level. *N* for left-handed *Ss* was too small to provide adequate data to test the effect; their mean choice was 10.42, with a standard error of 1.26. This is apparently in the wrong direction, and, though not significantly different from 9.00, nor from 11.18, is significantly different from 6.92 at the 0.05-level. The predicted match would be 6.92, using 2.18 as the size of the *BKD* as obtained from the right-handed *Ss*—if one assumes that left-handed *Ss* are just as left-handed as right-handed *Ss* are right-handed. The present result, although based on a very small *N*,

suggests that handedness is less pronounced in left-handed Ss than in right-handed Ss. In this connection, it should be noted that in the results of McPherson and Renfrew, the mean *BKD* of left-handed Ss was about half as large as the mean of right-handed Ss.

As for the kinesthetic figural after-effect, the average post-inspection match for all 60 Ss was 15.13, with a standard error of 0.38. The difference (4.01) between this value and the control value (11.12) was significant at far better than the 0.001-level; this result shows that over and above a significant *BKD* there also was a significant figural after-effect. The 51 right-handed Ss did not differ significantly from the 9 left-handed Ss in size of figural after-effect, with averages of 4.17 and 3.86 respectively. On the other hand, women averaged 3.08 and men 4.78, a difference significant at the 0.01-level. The women thus made higher control matches, and produced smaller increases over the control following inspection, than did the men; this suggests the possibility of a negative relation between *BKD* and figural after-effect. A test of this suggestion is provided by a correlation of control and post-inspection matches; but this correlation (-0.19 , Pearson product-moment) was not significantly different from zero at even the 0.10-level, and hence cannot account for the observed sex difference.

In conclusion, the present results confirm McPherson and Renfrew's findings, at least with regard to right-handed Ss; they also, however, separately show the kinesthetic figural after-effect. If the control measurements had not been made, the kinesthetic figural after-effect would have erroneously been considered an average distortion of about 27%; with the control measurements, the distortion, though still highly significant, becomes only about 13%. It therefore seems clear that appropriate control measurements of bilateral inhomogeneities in judged size should be made in future studies of kinesthetic figural after-effects, just as Brown has shown that such controls are necessary in measuring visual figural after-effects.

Wesleyan University

MICHAEL WERTHEIMER

SOME REMARKS CONCERNING THE USE OF THE METHOD OF ADJUSTMENT WITH IMPLICIT STANDARDS

In the classical and most familiar application of the method of adjustment, the observer causes some stimulus to vary until it is perceived as being equal to some standard or fixed stimulus. The standard stimulus has the following characteristics: (1) It is physically constant; (2) It is de-

scribed in identical terms as the variable, that is, it is in the same dimension of experience as the variable; (3) It is measurable independently of the measures taken on the variable. This physically defined, objective standard we shall call the explicit standard.

The paradigm of the method of adjustment may be used under conditions where the standard does not have these characteristics, that is, where (1) the standard is assumed, but not known, to be constant; (2) it is not typically measurable independently of the measures taken on the variable; (3) it is typically a non-tangible construct in a different dimension of experience than the variable. This kind of standard we shall call an implicit standard. It is the purpose of this paper to point out some ways in which ambiguities may arise from the use of the method of adjustment to an implicit standard, and to suggest a methodology which greatly reduces these ambiguities.

When a variable stimulus is adjusted to an explicit standard, as in the Galton bar experiment, certain measures may be taken which are: (1) the mean of the adjustments; (2) the difference between the mean of the adjustments and the standard (the so-called constant error) and (3) an index of the dispersion of adjustments around the mean adjustment.

In many applications of the method of adjustment to explicit standards, the constant and variable errors are themselves of prime interest. This application is usually in situations where questions of accuracy or precision of adjustment are studied. In many contemporary experiments, however, the constant errors of adjustment are employed as an index of the effect of some independent variable. In this case one is interested in the changes in the magnitude of the constant error of adjustment as a function of the particular experimental conditions that are imposed on the adjustment operation.

Suppose the constant error of adjustment is measured under two conditions, A and B . We could compare the difference between the constant errors under each condition by comparing the mean adjustments for each condition. In fact, this is the usual way the comparison is made. It is obvious that: mean adjustment (A) - mean adjustment (B) = constant error (A) - constant error (B) because constant error = mean adjustment - standard.

Thus, when comparing the differences between two constant errors the value of the standard need not be measured so long as it is certain that it has a fixed value. Only the mean adjustment under each condition need be obtained. This seemingly trivial fact assumes considerable importance when the method of adjustment to an implicit standard is analyzed.

A particularly neat example of the method of adjustment to an implicit standard is shown in the following study by Werner and Solomon.¹ In this experiment, *S* projected a negative after-image on a neutral disk. Concentric with this disk was a ring which could be varied in brightness. *S*'s task was to tell *E* how to vary the brightness of the ring until it appeared to be as bright as the brightness of the projected after-image. A schema of the results is as follows:

Experimental condition	Mean adjustment	Constant error
<i>A</i>	171.5	?
<i>B</i>	122.5	?

Since there is no measure of the standard independent of the measures on the variable, there can be no measure of the constant error of adjustment. The difference between the constant errors can be found, however, by subtracting one mean adjustment from the other, as was explained previously. If this procedure is followed, it must be assumed that the value of the implicit standard remains a constant, that is, that the value of the implicit standard *does not change* as a function of the experimental conditions. Werner and Solomon, however, make just the opposite assumption, that is, that the implicit standard (the brightness of the after-image) varies as a function of the experimental conditions. In this case the authors must assume that the constant error of adjustment is fixed, that is, that it does not vary as a function of the experimental conditions.

Clearly there is an ambiguity here. On the one hand, the difference between the mean adjustments could be 'explained' by stating that the brightness of the after-image did not change, but rather, that the accuracy of the adjustment of the ring to the brightness of the after-image was different under the two conditions. On the other hand, one could argue that the brightness of the after-image was different under the two conditions and this was reflected in the different mean adjustments.

Given a significant difference between the mean adjustments, the different inferences from the data are as follows: (1) The implicit standard is a constant, and the constant error of adjustment is varying as a function of the experimental conditions. (2) The constant error of adjustment is a constant, and the value of the implicit standard is varying as a function of the experimental conditions. (3) Both the implicit standard and the constant error of adjustment are varying as a function of the experimental conditions.

¹ Paul Solomon and Heinz Werner, Studies in contour: III. Negative after-images, this JOURNAL, 65, 1952, 67-74.

If the only measure taken is the mean adjustment, then the choice of interpretations must be made on the basis of theoretical plausibility and information obtained from other experiments rather than on the basis of characteristics in the data. Fortunately, with a slight modification of the methodology, the choice of alternatives may be so made, on a more rigorous basis, that the situation is much less equivocal.

Whenever it is desired to test the effect of different experimental conditions on the matching of an objective variable stimulus to an implicit standard, the measure taken should be the interval of equal judgments; thus the transitions 'greater than' to 'equal' and 'less than' to 'equal' are measured in each trial. This measure is completely analogous to the measure of the difference limen obtained by the method of limits. A comparison of the differences between these *DLs* will indicate whether the implicit standard is varying, whether the constant errors of adjustment are varying, or whether both are varying.

When using this measure it must be assumed that the best estimate of the implicit standard is the middle value of the range of equality, certainly a reasonable assumption.

If the range of equality is different in magnitude for each condition while the middle value of this range does not change, then one may infer that the implicit standard is not varying as a function of the experimental conditions but rather that the constant errors of adjustment are varying as a function of the experimental conditions.

If the magnitude of the range of equality is the same for each experimental condition, while the whole range has so shifted that the middle value of the range of equality is different for each condition, one may infer that the implicit standard is varying as a function of the conditions.

If both the magnitude of the range of equality and the middle value of the range of equality change as a function of the experimental conditions then it may be inferred that both the implicit standard and the constant errors of adjustment are varying as a function of the conditions.

This methodological procedure is especially useful when the effect of different starting values of the variable on the adjustment to an implicit standard is studied.

In the studies of egocentric space localization implicit standards are of great methodological importance. Some of these implicit standards are: (1) subjective vertical; (2) subjective median plane; (3) subjective horizon; (4) subjective body length. It is possible to find the range of equality for all of these standards, that is, all those values of the variable which are perceived as equal to the implicit standard.

In passing, it is worthwhile to note that here we are not interested primarily in the relation of the mean adjustments to some conventionally defined, objective standard. Since these conventional criteria seldom coincide consistently with the implicit standard (in the dark room, the plumb line vertical is not always perceived as being vertical) any question of precision or accuracy with regard to the conventional standard is completely trivial, if this is the only question posed. The more important question concerns the relation between the mean adjustments and the value of the implicit standard.

In conclusion, when the effect of different experimental conditions on the adjustment to an implicit standard is studied a range of equality (analogous to a *DL* by the method of limits) should be taken in each trial because (1) it gives the best estimate of the value of the implicit standard, and (2) it renders the interpretation less equivocal.

Clark University

THOMAS B. MULHOLLAND

ANOTHER PUZZLE-PICTURE

The accompanying figure is one to be added to the psychologist's library

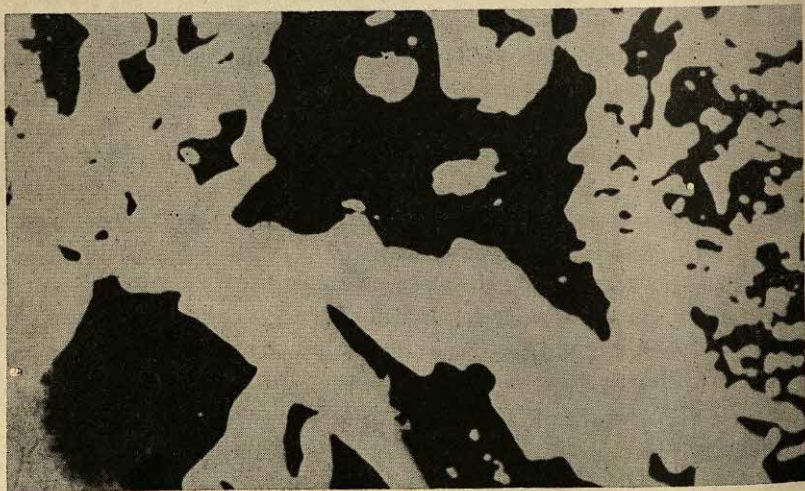


FIG. 1. FIND THE HIDDEN MAN

of puzzle-pictures.¹ The principle of concealment used in it is, as we believe, the same as in the hidden-cow picture described by Dallenbach.²

¹ My thanks are due Mrs. Dorothy Archbold for this picture.

² K. M. Dallenbach, A puzzle-picture with a new principle of concealment, this JOURNAL, 64, 1951, 431-433; More about the 'cow' puzzle-picture, *ibid.*, 65, 304-306.

Like the picture of the cow, this one makes no sense until it is organized as a full-face and shoulders, but this time of a human being. The upper margin of the picture cuts the brow, thus the top of the head is not shown. The point of the jaw, clean shaven and brightly illuminated, is just above the geometric center of the picture. A white mantle or serape covers the right shoulder. The right upper sleeve is exposed as the rather large black area at the lower left. The hair and beard are after the manner of a late medieval representation of Christ.

University of Utah

PAUL B. PORTER

THE XIV INTERNATIONAL CONGRESS OF PSYCHOLOGY

At the XIV International Congress of Psychology, which was held in Montreal June 7-12, 1954, the program of papers represented a striking departure from the practice usually followed at congresses and at most meetings of large associations. The committee in charge of the program made the bold decision to send out no call for papers. In an effort to achieve greater unity and coherence for the scientific sessions the members agreed that all reports would be by invitation only, and that speakers would be chosen on the basis of their ability to make significant contributions to the various topics and symposia which the committee had decided upon more than a year before the date of the Congress. The general plan was to have not more than two or three meetings in session at any one time, and to select for each meeting four or five speakers who would address themselves to the topic in hand. Dr. R. B. MacLeod undertook the task of arranging for the programs and speakers. The general opinion expressed by those in attendance was that the XIV International Congress set an excellent example of high standards of material presented and interest aroused. Discussions after each session were long and frequently heated.

The topics chosen for the symposia ranged widely over the whole field of modern psychology. The main emphasis was supposed to be on new and basic research and theory. Clinical papers, for example, pointed out the relation of diagnostic and testing procedures to fundamental concepts in abnormal psychology, while papers on industrial psychology presented important considerations for social psychology and for theories of personality. Abstracts of all the papers will be published in the proceedings of the Congress.

Meetings at which papers were presented took place in the morning and afternoon. Evening sessions were given over to special addresses. On Monday evening, June 7, Henri Piéron, President of the International

Union of Scientific Psychology, introduced the co-presidents of the Congress, Edward A. Bott and Edward C. Tolman. Bott spoke on the influences of organization on psychology as a science. Tolman dealt with the subject of performance vectors and the unconscious, during which he found time to digress impressively upon the intellectual repressions on both sides of the iron curtain. On Tuesday evening, June 8, Jean Piaget gave an address, *Structures perceptives et structures cognitives dans le développement de l'espace chez l'enfant*. A large audience on Wednesday evening listened to a lecture by Wilder G. Penfield, Professor of Neurology and Neurosurgery in McGill University, on the cortical record of the stream of consciousness. On Thursday evening Baron Albert Michotte van den Berck gave a detailed account of his views on perception and cognition with the aid of slides which demonstrated many of the experiments which he has performed over the years at Louvain.

Two special convocations were held on the closing day of the Congress. At the University of Montreal honorary degrees were conferred upon Herbert S. Langfeld, Albert Michotte, and Henri Piéron. Similar awards were made at McGill University to Edward A. Bott, Jean Piaget, and Edward C. Tolman.

At the XIII International Congress of Psychology held in Stockholm in 1951, the newly founded International Union of Scientific Psychology accepted the invitation of the Canadian Psychological Association to hold the 1954 Congress in Canada with the American Psychological Association as joint sponsor. The choice of Canada as the place of meeting seemed suitable for many reasons, not the least of which was the desire to spare foreign psychologists the difficulties which some of them might encounter in trying to arrange for travel in the United States.

The committee in charge of planning for the Congress was composed of six Canadian and six American psychologists, with Donald G. Marquis as chairman. Money was the main concern and headache of this committee, particularly after it became apparent that fees for membership were going to fall far short of expectations; but by skillful maneuvering on the part of several members of the committee, especially the chairman, the task of bringing the expensive Congress to a successful close without an appreciable deficit was happily accomplished. Donald O. Hebb was chairman of the Local Arrangements Committee, and to him and his co-workers the visiting psychologists from all over the world are indebted for the smoothness and efficiency with which housing, meals, transportation, entertainment, official³ receptions, and scores of other difficult matters, including the weather, were so nicely managed.

Registrations for the Congress were slow, and late. Final figures indicate a total registration of 876 full members and 114 associate members. These individuals came from 31 different countries, including the 4 from Russia who read papers on recent advances in conditioning. The relatively small number of Americans who attended the Congress was a source of surprise and disappointment.

At the meeting of the Assembly of the International Union of Scientific Psychology it was decided to hold the next Congress in 1957. Jean Piaget was elected President of the Union.

Princeton University

CARROLL C. PRATT

THIRTY-FOURTH ANNUAL MEETING OF THE WESTERN PSYCHOLOGICAL ASSOCIATION

The Western Psychological Association held its thirty-fourth annual meeting Thursday through Saturday, May 20-22, 1954, in the Wilton Hotel, Long Beach, California. The total registration was 840. One hundred forty-two papers, eight symposia, six special luncheons, an invited address, and a public information program were scheduled, in addition to the presidential address and the business session.

The presidential address, entitled "The Growth of Intelligence," was delivered by Dr. Nancy Bayley at the banquet Friday evening. Special luncheons were held by the Committee for the Scientific Study of Religion, the State Colleges Psychological Staffs, the Veterans Administration personnel, Psi Chi, the Society for Projective Techniques, and the Los Angeles Society of Clinical Psychologists in Private Practice.

A committee made up of representatives of UCLA, USC, Occidental College, Long Beach State College, and the Veterans Administration, under the chairmanship of James F. T. Bugental, was in charge of the arrangements.

At the business session Saturday morning it was decided that from now on the place of meeting is to be scheduled two years in advance. The president was asked to appoint a special committee on the structure of the organization to consider which of the informal procedures that have been followed in the past are no longer applicable now that the organization has reached its present size.

In the mail ballot subsequent to the meeting Allen Edwards was chosen President-Elect. The other officers, Neil Warren, President, Leona E. Tyler, Secretary, and George P. Horton, Treasurer, retain their present positions. As a result of the advisory mail ballot on meeting places, it was decided not

to hold a separate meeting in 1955 since the American Psychological Association is scheduled to meet in San Francisco, but to accept the invitation of the University of California at Berkeley for 1956.

University of Oregon

LEONA E. TYLER, *Secretary*

THIRTEENTH ANNUAL MEETING OF THE CANADIAN PSYCHOLOGICAL ASSOCIATION

The thirteenth annual meeting of the Canadian Psychological Association was held in Montreal at the Queen's Hotel on June 5-6, 1954. Total registration was 201. Eighteen papers, six symposia, and the annual business meeting comprised the program. Three special addresses were given during the meetings: Dr. R. B. McLeod presented the preliminary findings in his survey of Canadian psychology; Dr. Otto Klineberg discussed national stereotypes and their theoretical and national implications; and Dr. Julian Blackburn gave an account of his visit to the Inter-American Society of Psychology. Dr. D. C. Williams gave the presidential address.

The following officers were elected for the ensuing year: President Noël Mailloux, University of Montreal; President Elect, G. A. Ferguson, McGill University.

McGill University

DALBIR BINDRA

BOOK REVIEWS

Edited by M. E. BITTERMAN, University of California

Method and Theory in Experimental Psychology. By CHARLES E. OSGOOD. New York, Oxford University Press, 1953. Pp. vi, 800.

Psychologists have a way of getting ahead of themselves in the publication of research and waiting too long between summarizing and evaluative accounts of their progress. When we drop too far behind, it becomes impossible for non-specialists to keep up; worse yet, even specialists may feel stifled by accumulating data as they lose sight of implications and presuppositions.

Now Charles Osgood has brought a breath of fresh air to the central disciplines of experimental psychology. He has written a large and conscientious book. In it visual perception and learning have been treated with originality and thoughtfulness and with a kind of serious attention to theoretical issues that we have not had for some time. Other areas, psychophysics and sensory physiology, language and thinking, are almost as well handled. Osgood has really read the literature of these fields and has chosen intelligently among the critical facts and problems; for this is not a compendium—it is a text-book in the best sense of the word. Osgood is for the time being a teacher, and a most effective one. For one thing, he can write. You need never struggle with a sentence or a paragraph to know what it means. Besides, his writing is vigorous and interesting; it never intrudes between you and him. There is, of course, one thing even more important than an author's style—his willingness to think about materials that he presents from other sources. Osgood thinks. The clichés are absent; the pat answers rarely appear; the anti-theoretical attitudes that so often obscure real problems are unknown to Osgood.

Students and their teachers will read this book for a long time. Both author and publisher deserve our gratitude and praise for a volume which, because of its very great merit, requires extended discussion.

The first disappointment in the book is perhaps an inevitable one. There are many places where it is out of date. No book can stay up-to-date long, and efforts to make it exactly so on the day of publication are expensive. Still, this book seems to have been started about 1947, and the author's reading beyond that point is reasonably complete only on a few special issues that he has pursued into the 1950's. No one could fail to be struck by the lost opportunities that result. Some scattered examples will illustrate what I mean. In discussing retinal bio-chemistry, Osgood states that "The regeneration of visual purple from its own end products must require the addition of energy from some external source," yet Wald and Brown published their demonstration to the contrary as early as 1950. In a detailed treatment of the possibility of understanding figural after-effects in terms of Marshall-Talbot concepts, Osgood remarks that these effects "do not require the postulation of novel electrical fields in the brain." Yet such fields have long passed the stage of postulation; they have been discovered and recorded and experimented with; and Köhler and Held had already published some of their observations in 1949. A generally thoughtful treatment of experimental analogues of repression omits Zeller's papers of 1950 which

come closer to meeting the necessary criteria than most. Harlow's 1949 work on learning sets is not mentioned anywhere although it is centrally relevant to several issues that occupy Osgood. It is regrettable that these and other discussions could not have been modified in the light of later evidence. Each reader will find his own specially poignant examples.

Sensory processes. The early sections of the book, on sensory problems, are well done. We have had a number of good, high-level treatments of most of these problems in recent years, and Osgood's presentation compares favorably without attempting to replace more specialized treatments. In a few places, notably in his discussion of temperature sensitivity, he is especially helpful because of his clarity and the originality of his arrangement of arguments. For the rest, he is traditional in his approach; sound, careful, and adequate in representing opposing points of view.

Perceptual processes. It is in the section on visual perception that Osgood first begins to impress his own psychological conceptions on the reader. It is impossible to read this section without admiration; it is equally impossible to read it without queries, criticisms, and occasional outrage. I think it is fair to say that no contemporary author has written so fully and so clearly about the central problems of perception. We have been especially unfortunate in recent secondary accounts of these topics, and Osgood's is the best general treatment since Woodworth's in 1938. In most ways it is better than Woodworth. Nevertheless, one finds confusions and contradictions that ought to be rationalized or amended in another edition.

I pass over with a wince Osgood's identification of isomorphism with mind-body parallelism; perhaps more careful reading of Köhler or more philosophical sophistication will someday rid psychologists of this notion. But there are more serious confusions in the same area. For instance, Osgood quotes with apparent approval the strange contention of K. R. Smith that the existence of stroboscopic motion between hemispheres (a fact that had long been familiar, by the way) creates difficulties for Gestalt theory. In fact, a field theory of some kind seems to be the only view capable of dealing with this phenomenon at all. Exactly the same issue arises in the discussion of Werner's theory of stereoscopic vision, where Osgood's dependence on data from strychninization leads him to reject a hypothesis that is in fact less damaged by hemispherical separation than is any neuronal connection theory—even if the strychninization studies are really relevant.

Osgood is less easily satisfied with 'unconscious inference' as an explanation of psychological facts than many contemporary writers; yet he falls back on such concepts with disturbing regularity when the going gets rough. Unfortunately his criteria for distinguishing between perception and secondary process are not consistent. Here is an example: in the account of brightness constancy, Osgood remarks, "The camera doesn't *perceive*," and he goes on to show that the camera records a mis-match where *O* thinks two disks are alike. No one can quarrel with that. In his discussions of the Ames demonstrations, however, Osgood presents as perceptual problems situations in which the camera is 'fooled' as much as the human eye. He even attributes the human illusion (though presumably not the photographic one) to past experience.

A kind of Helmholtzian view is especially apparent in discussing movement and the constancies. In the former case, Osgood falls back on the assumption that identification (of a second exposure with the first) is necessary for stroboscopic

motion, concluding then that identification is prior and permits an inferential process. The facts are otherwise. Two objects that would never in this world be identified under other conditions will give rise to movement under stroboscopic conditions, maintaining a kind of identity, to be sure, but changing shape, size, and color in the process. Identification is a psychological problem in its own right; and besides, the conditions for movement change the conditions for identification.

Apparently Osgood is prepared to explain brightness constancy in fish as a matter of past experience, though the fact that some flat-fish change their skin color to match ocean bottoms in almost complete accord with constancy leads me to wonder what kind of learning would explain the acquisition of such a skill. Osgood shows here and elsewhere a real resistance to the proposal that animal nervous systems may react to relationships rather than to punctate stimuli; it is hard to reconcile this resistance with Osgood's elegant demonstrations in other parts of the book that contemporary concepts of nerve conduction are capable of handling complex relational matters. The stubborn attempts to force all perception into the mold of piecemeal learning persists even into little things: Osgood says of children's drawings that the child draws things as he *knows* them to be; circumventing the obvious fact that the child draws them as he *sees* them, and ignoring the problem of how the child discovers which of many appearances he is to use as the "known" one.

Related to the inference interpretations is a recurring tendency to slip out of stimulus-concepts into poorly defined conceptions of seeing or knowing. For instance, in criticizing Wallach's hypotheses about achromatic interaction, Osgood denies that the immediate surround is the only carrier of information about illumination; he says that if a blacker surface used as a background is "obviously under stronger illumination" constancy effects are still found. He does not indicate what the stimulating conditions might be that would make a darker surface obviously more highly illuminated; it is a nice problem. These things are troublesome just because Osgood is so careful and so provocative in most of his discussion. Just when you have begun to think that he has the problems of constancy well in hand, you may be brought up short by such a lapse.

There are a number of individual curiosities about the visual portions of the book. It may be worth listing a few of them. Nowhere in the discussion of the rôle of complementaries in color theory does Osgood mention the fact that mixture-complementaries and after-image-complementaries are not identical. He places a kind of reliance on the alleged facts about learning to see after cataract operations that is difficult to understand in an author who has read the mixed-up and inconclusive reports of Senden. Despite his association with the Dartmouth group and his obvious sympathy for the Amesian point of view, Osgood does not refer to the work of Ogle on binocular mechanisms—surely the finest work to come out of the Dartmouth Eye Institute.

Learning. In the sections on learning, Osgood has made a more than valiant effort to treat all the issues of learning theory in considerable detail. His erudition is again fully equal to the task, and he again shows his penetration in questions, suggestions, and reconstructions. Students will find this part of the book a useful supplement to the other fine secondary sources already available. It is encyclopedic enough for anything except exhaustive bibliographic search, and it has a kind of theoretical unity that is hard to match elsewhere.

Unfortunately the number of interconnecting issues in present-day learning theory is very great; modifying a hypothesis to fit new data creates a hazard of making the hypothesis incapable of explaining something else that it formerly handled. Osgood has a way of not letting the left hand know what the right hand is doing, so that he solves one problem with an interpretation that would not be permissible in another section of the book.

For example, in discussing the phenomena of extinction, Osgood makes it clear that the response-competition theory founders on data about the effects of drugs, about spacing of trials, and so on. Yet when he finally comes to a summary interpretation of response-decrement, all that seems to be forgotten, and the learning of competing responses is enthroned as a major principle. The treatment of conditioned inhibition is equally inconsistent; sometimes it is acceptable, sometimes it is not. In much the same way we get generalization cut to fit the problem under attack. In discussing discriminative learning, Osgood introduces without reservation the notion that generalization increases as learning increases, although he shows much later that he is well aware of such work as that of Gagne which requires a more complicated relationship between degree of learning and degree of generalization.

On the issue of continuity, Osgood is traditional and clear—he demolishes Lashley's position with due courtesy and sympathy, and without hedging—but his much later discussion of transfer and retroaction is written as if completely insulated from the continuity problem. Studies of Underwood and of Gagne and Foster (and of others not mentioned by Osgood) demonstrate that a strict continuity position is untenable. If it were true that every trial increases response-strength, there could be no reversal in the amount of transfer as a function of practice. It is ironic that Osgood should quote with approval Underwood's hypothesis that some form of discrimination supervenes, since it is exactly discrimination that was supposed to be explained as a continuous process a few pages back. We cannot really hold Osgood accountable in any serious way for these gyrations. They exist because of the state of our knowledge in mid-twentieth century and because we try so hard to find general postulates and principles to explain the enormous amount of empirical data classified under the heading of learning. Nevertheless, the reader should beware of reading separate sections of this long portion of the book lest he miss important evidence that turns up only under another argument.

Two theoretical contributions of Osgood's require special attention. The first is his mediation hypothesis which adds to the formal equipment of Hullian theory a concept often suggested in slightly different form by others. Osgood's mediator is much like Hull's 'pure stimulus acts' and Guthrie's 'response-produced stimuli.' I find it indistinguishable from Seward's 'surrogate response.' Like Seward's intervening variable, Osgood's has both stimulus and response properties; it is in fact a stimulus-producing response to overt stimuli. Added to Hull's general system and treated as if it obeyed the laws of that system as they apply both to stimuli and to responses, it increases the power of the theory considerably. It makes possible conditioning-explanations of latent learning, place learning, and other phenomena that cognitive theorists have claimed to spell doom for associationism. We may reasonably ask, however, why 'mediating response' at all? Once we have passed beyond naïve behaviorism to a language of constructs what is so sacred about responses? Why not call these hypothesized events *ideas*?—or, more specifically, beliefs, ex-

pectations, memories and other varieties of *idea*? It is a strange fact that the only trace of dualism extant in scientific psychology should be found in the modern behaviorist who appears to be frightened by certain terms because they seem to him to describe a different realm, an incommensurable substance, from the ones with which he is used to dealing. Or, if mental terms are too strong medicine, why not talk of a central nervous process instead of a response? The constant preoccupation with response suggests the old motor theories of thought with all their inconsistencies and confusions; if we do not mean response in any real sense, why not talk of the (presumably) cortical event that may or may not lead to response in a given instance but is certainly available for association whether it does or not?

My second problem is Osgood's treatment of similarity in retroaction and transfer. Here he repeats essentially what he wrote a number of years ago in a journal article. It goes like this: (1) Similiarity can be a matter either of stimulus or of response; (2) in either case, it can be defined operationally by reference to functional identity in the laboratory situation; (3) to the extent that two learning situations have functionally identical stimuli, transfer will tend to occur, zero transfer occurring only when the stimuli are unrelated; (4) given a tendency to transfer, the amount and direction will depend on the similarity of the responses required in the two kinds of learning—functional identity of responses leading to a maximal positive transfer, intermediate degrees leading to progressively negative transfer, and antagonistic responses leading to maximum negative transfer. These four "empirical laws" are then combined to produce a three-dimensional graph showing "the surface of retroaction and transfer," for what is true for transfer is also said to be true of retroaction.

There are too many difficulties with this view to permit exhaustive treatment here, but let me point out a few inadequacies. The similarity dimension is not really defined. It is only anchored. Granting that we can agree on the meaning of "functional identity," where do we go from there? The implication is that as physical aspects of the stimulus change, similarity (and hence transfer) changes monotonically, but the allegation that this is not so has always been at the heart of the similarity controversy. In another lapse of memory, Osgood neglects here the issue of octave-generalization that he raises in another place. If a response of the kind learned to a tone of 1000 cps. can be more readily produced by a new tone of 2000 cps. than by one of 1500 cps., what happens to the operational definition of similarity?

Moreover, a transfer dimension that must fit both memory and performance creates difficulty and confusion. The difference between positive and negative (performance) transfer is an extrinsic one: two rats with the same training will perform alike in a new maze (both, say, turning to the right because that was what they did before); if it is *E*'s whim to reward one for this response and not the other, we score two identical performances differently, one as positive transfer and the other as negative. Retroactive interference, on the other hand, produces what appears to be a failure of the present stimulus to arouse a particular past association. We might be able to make a rat forget which way to turn on Maze *A* by interpolating a lot of other mazes. If we do, it will prevent both positive and negative transfer from Maze *A* to the present one. In other words, the identification of positive transfer with retroactive facilitation leads to difficulties. Osgood's treatment seems to imply that communication between a present process and an old association must lead to

repetition of the originally learned act. That is manifestly not so. A high degree of similarity between the present stimulus and an old one may lead to *recall* of the old situation without leading to the response that then occurred. Recall itself may even lead specifically to inhibition of the response. I see no hope for clarifying the transfer-retroaction problem until we learn to distinguish between riding a bicycle and thinking about riding one, until we give up this outmoded preoccupation with response and substitute an equally serious concern for central processes that make responses possible. I have a hunch that mediation theory is perfectly capable of handling these distinctions, and I hope Osgood will someday attempt to work them out in detail.

Here, as in the section on visual perception, there are *curiosa* which I feel constrained to sample. Osgood says in one place that the randomness of early stages of instrumental learning is a concept fostered by gestalt critics; this statement is so completely at variance with history that I can only suppose there is some ellipsis or printing error involved. Thorndike is the father of the random nature of the cat's behavior, and the "gestalt critics" were most vigorous in their opposition to the whole notion. Another strange item is the statement that "higher order conditioning is certainly greatly extended in human learning." Attempts to prove this contention have failed signally, and one can only wonder where Osgood's certainty comes from. Again, the author says with conviction that all theorists agree that contiguity (between stimulus and response) must occur before learning can take place. Certainly not Tolman; certainly not Köhler; certainly not Katona—to name only a few men whose work Osgood quotes in other places. What can such a sentence mean? These are blemishes on an otherwise excellent section on learning, a section to which I expect to return again and again in connection with both teaching and research.

Symbolic processes. This last portion of the book contains much interesting material not summarized elsewhere, but it is not up to the earlier portions in effectiveness of organization or in incisiveness of writing. Perhaps that is the fault of factual paucity and not the fault of the author.

In these pages Osgood several times shows signs of recognizing the close relationship between his mediating reactions and what we normally call ideas, and in one or two places he even seems prepared to admit that they need not be responses at all. At other times he persists in what seem to be gratuitous efforts to force language and thinking into a behavioristic mold. This kind of thing is particularly evident when he writes of the mediator attached to a word as "being some fractional part of the total behavior elicited by the object." In other places we find vacillation: Osgood criticizes C. W. Morris because of "the fundamental danger that 'dispositions' may serve as mere surrogates for 'ideas' in this system," but then two pages later he writes, "the representational mediation process is a hypothetical construct . . . it could be entirely cortical without invalidating behavioral implications."

There are some fine things in this part of the book. Osgood's own explorations of the problem of measuring meaning are extremely suggestive. His comments on other people's work is often productive and insightful: witness his suggestion for reversing the order of training in Maier's 'reasoning' experiments. A real achievement is the imposition of some unity upon matters as diverse as concept-formation, the statistical structure of language, and synaesthetic thinking.

The critical reader will find a number of surprises. The defense of associationism

leads to certain kinds of carelessness. In discussing Birch's studies of learning in the chimpanzee, for instance, Osgood concludes, that "past experience played a rôle in the insightful solution of problems," but this does not mean what seems to be implied. Particular S-R sequences had not been used before, and in the phrase "past experience" lurks the entire controversy about how experience works. Osgood seems to be saying that Birch's studies undermine the notion of insight and strengthen an association theory, while in fact the issue is left in doubt. Or again: in accounting for the results of Maier's experiments on human learning, Osgood writes that "dispositions toward the various skill sequences . . . become associated with such perceived properties." How does a disposition become associated? The question at issue seems to be begged. No attempt is made to show that dispositions are really responses for another sixty pages, and even then it is not done convincingly; but here in one phrase the real problem is by-passed. A final example of the same kind is Osgood's dismissal of unlearned acts from the category of sign-behavior, thereby begging at least part of that question and introducing an *a priori* limitation of an historical kind into what should be a functional definition. Surely the question of whether or not a particular act is a sign-process must be decided in terms of the contemporary properties of the act itself in relation to total behavior; not on past history.

The minor curiosities of this last section include the following sentence referring to an experiment of Foley and Cofer: "Although recall after reinforcement with synonym list II was not significantly superior to that after the control list, it is surprising that there should be any difference at all. . . ." What difference? The failure of the significance test means precisely that, within the error of measurement, there is no difference.

In summary, it is important to say again what a fine and stimulating work Osgood has turned out. It has a kind of integrity that one finds all too seldom in psychological writing. Every page attests to the author's scholarship, his thoughtfulness, and his understanding. Nowhere does the reader feel that a topic is included for extraneous reasons or a problem skimmed in the fashion of mechanical library research. Osgood's own contributions to psychological theory will make this book useful for a long time to come by providing meat for experimental chewing in many laboratories. In addition, he has provided a textbook and reference work that has no contemporary peer and may well prove the 'bible' of graduate students and advanced undergraduates for some years. He and his publisher are to be congratulated.

Swarthmore College

W. C. H. PRENTICE.

Introduction to Psychology. By ERNEST R. HILGARD. New York, Harcourt, Brace and Company, 1953. Pp. x, 659.

One of the rights of the reviewer of an introductory text-book is to wander afield and to express his views on the philosophy of science, the state of psychology, and the CRISIS FACING US. Because I think that Hilgard's new text is a good one, and because Hilgard knows the high regard in which he is held by his colleagues (including me), I feel certain that I can avail myself of that right (at his expense, and to some extent irrelevantly) without risk of misinterpretation.

The institutions, professors, and students who annually become enmeshed in

introductory psychology courses are large in number and varied in kind. Encouraged by the variability, every instructor in psychology carries a 'new and different' textbook manuscript in his briefcase; and, in view of the tremendous market created by so many students, every publisher seeks to ferret out these manuscripts and add them to his list. Of the writing and publishing of introductory textbooks, therefore, there is no end—nor is there likely to be. These comments are not intended critically—at least not completely so. I feel that it is good that so many students are subjected to some education in psychology; I feel that different groups of students require textbooks written at different levels of complexity and thoroughness; and, finally, I feel that psychologists should be encouraged to produce varying textbooks. These are all good and sufficient reasons for suffering many different textbooks to come unto us. Our field is simply not well-enough structured (in terms of either facts or notions) to permit a 'standard' textbook even for the introductory course. Nevertheless, once having established myself as four-square for the virtue of tolerance and as yielding to no man in my abhorrence of the sin of dogmatism, I feel it necessary to specify in more detail what I mean by virtuous tolerance and sinful dogmatism. In the light of these specifications I shall then discuss Hilgard's new text.

To say that different groups of students need textbooks written at different levels is not to justify two current fads which seem deliberately and cynically to underestimate both our students and our colleagues. The first of these is the assumption that textbooks for college and university students should be written with as low a 'Flesch-count' as possible. It is my impression that a number of our recent texts have deliberately sacrificed a competent statement of a complex idea on the altars of the one-syllable word and the uncomplicated sentence. Just as 'Gobbledygook' is an abomination, so is 'Simplicideblah.' It is in the nature of the beast which is science that some of its ideas and concepts *are* complex, abstract, and technical. Now, it is probably true that even a complex and abstract concept may be expressed in one-syllable words if you use enough words—but why do so? It *may* make it easier for the reader to follow, but it may also make that much less graceful a reading, that much more boring a reading, and that much less rewarding a reading. Further, it will not prepare the student for advanced reading or work in the field. Any science, indeed any discipline, demands its specific language. Any textbook which presumes to introduce a student to the science of psychology can hardly avoid the use of technical, polysyllabic, complex, and abstract expressions. My tolerance for discussions at different levels of difficulty does not, then, extend to textbooks written with the going vocabulary of the high-school freshman in mind.

The second fad is based on the assumption that the more closely a textbook simulates a current copy of *Life Magazine*, the easier it becomes for the student (and the professor, I assume) to become ego-involved in such concepts as S-R theories of learning, retroactive inhibition, interaction-process-analysis, and sign-gestalt readiness. It is my impression that a number of our recent textbooks have irrelevantly (and irreverently) over-seasoned and over-peppered discussions of psychological concepts with 'text-book cheesecake'—pictures, photographs, and cartoons that add not a whit or a smidgin to such discussions. I am not here referring to pictures of apparatus or even to photographs of eminent psychologists. I am referring to such over-worked 'corn' as a picture of several three-year olds

playing together in a sand-pile which is captioned: "Social inter-action." I suspect that for this kind of childishness (which adorns so many of our text-books) the publishers, more than the authors, are at fault. It is perhaps another instance of what I might call the 'Law of Hollywood Contagion': "IF ONE STUDIO MANUFACTURES A SUCCESSFUL HORSE OPERA THEN ALL OTHER STUDIOS MUST MANUFACTURE HORSE OPERAS." This is known as 'giving the public what it wants' without any regard for the notion that the consummatory behavior of the public is, among other things, a function of the choices available to it, and that the public may prefer, and even consume avidly, *other* products if only those other products are made available. And so, one or two publishers having successfully sold hundreds of thousands of copies of introductory textbooks replete with second-hand 'cute' or 'interesting' pictures, *every* publisher then seeks to add such books to his list. I assume that the next step will come when some enterprising publisher puts out an introductory text complete with 'pop-ups' and pictures to be colored by the students and professors. I would suggest (tentatively and hesitatingly) that it is just conceivable that a well-written, serious, and intellectually exciting textbook *without* such pictures might turn out to be a financially successful publishing venture—and at a considerable savings to the student who, in the last analysis, must pay for the slick-paper reproduction of these warmed-over pictures. My tolerance for textbooks of different levels of complexity or thoroughness does not, then, extend to picture-books accompanied by some textual material.

I have said that our field isn't structured well-enough to permit us to decide upon one 'standard' textbook, but psychology is not completely without form and without shape. Actually psychology is in very good supply of orientations, theories, and systems, and one would think that such wealth would result in a many textbooks written in terms of some orientation, theory, or system. This was true at one time, but not so now. *Someone* is laboring under the impression that students do not want theoretical systematization, that our colleagues are either unappreciative of the need for such systematization or else incapable of teaching a theoretically coherent course, and that the less integration the better. Now it is obvious that any *one* textbook, written from one systematic point of view, will appeal only to a limited market. But that is as it should be! Let there be many textbooks, each one clearly and unashamedly designed to present one systematic ordering of the fundamental data of psychology. The battle-cry should be not "How can I manage to break X's strangle-hold on the market?" but "How can I best write an integrated presentation of psychological data for those instructors who sympathize with my approach." This contention may only display my naïve business-sense. I don't think so. So-called 'business sense' has nothing to do with the problem. No textbook should be written unless it is designed to advance psychology as a science, or a profession, or a practice, and I am committed to the notion that psychology will not so advance unless we persistently strive to seek *understanding* of our data and to convey that understandings to our students. Felix Bloch, co-winner of the Nobel Prize in Physics for 1952, began his acceptance address with the remark: "It is a tribute to the inherent harmony and organic growth of our branch of science that every advance in physics is largely due to the developments that preceded it." Now I know that psychology has not yet achieved that degree of "inherent harmony and organic growth" which Bloch ascribes to physics, but, at the

risk of sounding wistful, is it too much to ask that at the very least the various chapters in an introductory textbook be *somehow* interrelated, and *somehow* connected, and *somehow* harmonious? Must each chapter be a completely closed *Psychological Bulletin* article (written at a high-school level) having nothing to do with the chapter preceding or the chapter following? Is this really the way to teach a science?

I have already said that I think Hilgard's new text is a good book, but it is, to me, somewhat disappointing. I had expected much more. Of the sins I have outlined above, Hilgard's book is guilty on one minor and one major count. In the first place, the 'packaging' of the book is in the best style of the 'overdone' school. I have the results of some research on this matter. According to my count the student who pays \$5.75 for this textbook is paying for about 40 pages of useless picture-book fillers—ranging all the way from a full-page picture of an innocent babe at sleep (with the caption: "Most of infant behavior has a physiological basis") to a half-page picture of women garment-workers at the work-bench. Perhaps Hilgard's text is not the worst sinner in this respect, but there is enough *fol-de-rol* to distress me and to cause me to resent the publisher's evaluation of American universities and psychology professors. This whole matter may, however, be primarily one of taste, and perhaps the sin should be regarded only as a minor one.

Of more consequence is the failure of systematization. Here I am tempted to say that, as far as inherent harmony is concerned, Hilgard's book has twenty-five chapters, but that would not be completely fair. Instead I might more fairly say that Hilgard has several relatively independent books within one cover. He has sections on "Growth and Development," "Motivation, Emotion, and Adjustment," "Learning, Perception, and Thinking," "Individuality," "Psychology Applied to Personal and Social Problems," and "Psychology as a Science and as a Profession." Certainly it would be difficult to regard this textbook as 'academic' or narrow—it covers much of the content of what is called 'Psychology' today—but nowhere is Hilgard's well-known eclecticism more obvious. Each section calls upon a different set of hypotheses and principles. It is a bit puzzling to discover, for example, that chapt. 9 ("Individual Modes of Adjustment") has no reference to chapt. 10 ("The Nature of Learning"). Actually one wonders why the student is introduced to problems of adjustment before he has had an opportunity to learn anything about learning, forgetting, perception, and thinking; but, even after the student has been exposed to the material on perception and learning, no use is made of it in the chapter on "Mental Health and Readjustment Techniques" where one would suppose that principles of *learning*, at the very least, would have some relevance. The reason for the failure to refer to these principles is especially obscure since in the introduction to the latter section Hilgard points out that "Wise application of the principles of psychology can furnish help in guiding individual lives, in managing human relations, and in predicting the direction of social change." Aren't principles of learning, perception, and thinking "principles of psychology" or are they principles which cannot be put to "wise application?" Of course, a serious attempt to apply the principles of learning, perception, and thinking to psychodrama, client-centered counseling, group-therapy, psychoanalytic transference, abreaction, and so forth (some of the topics dealt with in chapt. 19) would have required a great deal of

theoretical thinking-through as well as theoretical courage, but anyone who would add another book to the long list of books we already have should be willing to do just that. "Integrate, integrate, integrate" is the shibboleth of the day. But who is to do the integrating? The student? Is he supposed to integrate psychoanalytic notions with S-R notions, with 'field' notions, with 'maturation' notions—and with a dozen other *unrelated* notions to which Hilgard has recourse as he goes from one section of the book to another? It should be clear that I am not objecting to the wide coverage of the book, but I should like to have seen more organic connections on this water-front. The wide coverage, indeed, is one of the strengths of the book, although even here I have a few minor questions to raise. Some sections have dubious value for an introductory book in psychology—such, for example, as the discussion of the number of jobs available in psychology today and what psychologists earn; some sections which should have been included were not—for example, material on basic neurology and statistics.

On the positive side of the ledger, several points should be made. In the first place, Hilgard is *not* guilty of the first sin I have discussed in my opening remarks. His writing is good. He is not afraid of a complex idea or a technical expression. He does not write down, nor is he unnecessarily and boringly repetitious. You do get the feeling that he had a University student in mind when writing this book. He avoids simplicideblah without falling into the trap of gobbledygook. His 15-page glossary of technical terms and "common words when used in psychology with restricted meanings" is well-done and will be appreciated by both the student and the instructor. From the point of view of style, the book is a difficult job well done. I have already referred to the wide coverage of material. The wealth of experimental data, clinical material, charts, and tables compares favorably with any current textbook in introductory psychology. *Within* any one chapter all of this material is efficiently and intelligently organized, and adequately integrated. In many of the chapters Hilgard attempts to summarize and draw out implications and general principles, which he does most consistently for the chapter on "The Nature of Learning." His level of aspiration here is good, but I'm not too certain of his level of achievement. *Item*: a "general principle of associative formation" is stated on p. 217 as follows: "*associative strength between two items is increased by repetition provided the items occur together and are similar or vivid or meet other conditions favoring associative connection.*" The last clause of this principle—starting with the phrase "*or meet other conditions*" is, in my opinion, one of the most amusing bits of circularity I have seen in psychology. Hilgard nodded at this point. *Item*: a principle "to account for recall guided by set" again emphasises how circular we can get when we try to *explain* (isn't that what Hilgard means by "to account for") stimulus-response correlations by recourse to stimulus-response constructs: "*In associative recall under conditions of specific set, those items tend to be recalled that fit the specifications*" (p. 218). Does this principle "account for recall guided by set"? But of course we should not rail against Hilgard—he is illustrating here what so much of psychology is today.

Until someone writes my introductory psychology text, Hilgard's will do, and given what *is* available to the public today, I recommend it heartily to all teachers of all persuasions.

University of California

DAVID KRECH

Advanced Statistical Methods in Biometric Research, by C. RADHAKRISHNA RAO. New York, John Wiley and Sons, 1952. Pp. xvii, 390.

The author's stated object is to present a number of statistical techniques of relevance to biological research, particularly those dealing with multivariate analysis and classificatory problems, topics which are usually passed over lightly or omitted in the standard textbooks on statistical inference and experimental design. Following a highly condensed review of the algebra of vector spaces, matrices, and determinants, Rao takes up the binomial, poisson, normal, gamma, beta, and the multivariate distributions, illustrating throughout the fundamental importance of least-squares in distribution theory. The third chapter deals first with applications of the least-square technique in the estimation of parameters, proceeds to the usual tests of linear hypotheses, as in analysis of variance, and ends with the theory of statistical regression. A thorough, up-to-date presentation of the general theory of estimation—concepts of consistency, information, sufficiency, and efficiency—is to be found in the fourth chapter, which emphasizes the method of maximum likelihood. In the next two chapters Rao discusses problems of specification and tests of homogeneity, largely employing the chi-square distribution.

The remainder of the book is devoted to multivariate analysis and the utilization of multiple measurements in problems of classification. This section should prove of special interest to psychologists, largely because most of the material is new or has hitherto been relatively inaccessible. Extension of univariate analysis of variance to the case of multiple variates has been slow in spite of great theoretical interest. Rao has attempted with fair success to develop a unified approach to the problem of tests of significance in multivariate analysis by developing the concept of analysis of dispersion and applying it with several concrete examples. Briefly, analysis of dispersion is the technique of analyzing the variances and covariances of multiple correlated variables into various categories with the corresponding distribution of degrees of freedom. Comparison of the mean product matrix due to deviation from the null hypothesis and the mean error matrix is accomplished in terms of criteria developed from an extension of the method of discriminant function analysis. The method appears to be a natural extension of the comparison of variances in univariate analysis of variance. Consequently, it should find valuable application to psychological problems in which a generalized test of significance for k samples and p variables is desired.

Of particular interest to many will be the thorough and original discussion of statistical inference in classificatory problems. Following a brief introduction to the theory of Neyman and Pearson dealing with test-power, Rao takes up Fisher's discriminant function and the difficulties encountered in its application. Several modifications are suggested and developed which are of great practical significance for certain problems of classification.

The last chapter presents a theory for measuring the distance between two or more populations and the determination of group clusters or constellations. This discussion is particularly pertinent to the problem of classifying different groups of individuals on the basis of a pattern-analysis, a problem about which psychologists have recently become increasingly concerned. Out of this theory is developed a method for representation of p -dimensional data in a smaller number of dimensions with a corresponding test of significance for the residual variation. The objectives are

markedly similar to the more widely known methods of factor analysis although the approach is distinctly different.

Examples of application are given in detail throughout the book, making it considerably easier for the reader with only elementary training in mathematics to see the potential usefulness of the statistical methods presented. Almost all illustrative data have been drawn from physical anthropology, and the reader will have little difficulty in visualizing the concrete problems. Two examples, both drawn from the same study, deal with psychological data. One problem is concerned with determining how scores on three psychological tests can be used with a minimum of error to classify army recruits as normal or in one of five neurotic categories. The other deals with a test for reduction in the number of dimensions from six to two for classification according to neurotic condition.

Written for the advanced student of theoretical statistics and biometric methods, this is a book which few psychologists will be able to understand fully. Although it deals with problems directly relevant to psychological research, the presentation is terse and highly mathematical.

WAYNE H. HOLTZMAN

Symposium on Fatigue. W. F. FLOYD and A. T. WELFORD, Editors. London, H. K. Lewis and Company, 1953. Pp. viii, 196.

The twenty papers brought together in this volume were presented originally at the Symposium on Fatigue sponsored by the Ergonomics Research Society in March, 1952, at the College of Aeronautics, Cranfield, England. The opening paper, by Sir Frederick Bartlett, attempts to establish a "set" for the members of the symposium by defining fatigue broadly as "a term used to cover all those determinable changes in the expression of an activity which can be traced to the continuing exercise of that activity and which can be shown to lead. . . to results within the activity which are not wanted" (p. 1). This general set appears to have been adopted wholeheartedly by the participants, for here is a collection of papers whose diversity in method, in theory, and in purpose, suggests that the current state of research on fatigue consists mainly of scattered, many-sided, and often independent experimental attacks on isolated and frequently unrelated parts of a little understood problem.

This great diversity may be illustrated in terms of a few specific approaches. Welford, for example, takes the position on the problem of measurement that the "main phenomena" of fatigue have been exposed and that they may be subjected to classification in terms of their effects, such as decrement in performance (pp. 190-191). On the other hand, R. C. Browne and T. A. Ryan, in separate papers, attempt to show that fatigue may perhaps be better understood in terms of concepts other than fatigue. Browne asserts that prolonged performance may result in subjective symptoms and change in performance, but he argues that fatigue is not a single entity whose measurement will result from continued research on measurement. Ryan, in turn, re-emphasizes much of what has appeared in his recent work on *effort*, defined as an *experience* and measured by physiological indicators which are correlated with the experience. Other investigators, among whom are H. C. Weston, H. R. Noltie, Erling Asmussen, and E. H. Christensen, appear to emphasize the practical problem of outlining the conditions which contribute to fatigue, with no attempt to define fatigue precisely.

A number of papers deal with the interesting and difficult problem of measurement

and identification of fatigue-states. One paper, by J. B. Eagles, A. M. Halliday, and J. W. T. Redfearn, reports a series of experiments which show that rate and amplitude of finger-tremor are correlated with exhaustion and acute exertion but not with chronic exertion. In another paper, J. L. Kennedy points out some practical problems in the employment of the recently developed 'Alertness Indicator.' Kennedy believes that, in spite of good correlations between muscle potentials and performance in laboratory tasks, no really sensitive measure of 'alertness' is yet available for practical use. Other papers by R. S. Schwab, by D. E. Broadbent, and by W. T. Singleton, describe attempts to control the motivational factors which contaminate the measurement of muscular fatigue (Schwab) and to classify measurable aspects of fatigue in terms of breakdown of the focus of attention (Broadbent) and reduction in output or in accuracy of performance (Singleton).

This collection of papers suggests to the reviewer two important trends in the study of fatigue. The first is a trend toward diversity in definition, with major emphasis on operational concepts (such as performance) which have received much emphasis in other fields, like learning. The other trend is a broadening of the scope of measurement, as shown by attempts to identify physiological states, either neural or muscular, which are correlated with the experience of fatigue or with changes in performance. The latter trend seems to result from a broader acceptance of the possibility that fatigue, resulting from continuous performance, may occur long before any real decrement becomes manifest. Future research may uncover methods for measuring neural or other bodily states which may in turn provide a more common ground for experimentation. Until such methods are found, further progress on the problem of fatigue will probably continue to consist largely in the isolation of specific factors within specific tasks that lead to specific changes in the performance of those tasks.

In this volume there is a wealth of material to stimulate the thinking of the most critical reader. A number of new techniques of measurement are presented and several interesting lines of approach are outlined. The symposium will be of interest to the reader who is looking for a summary of recent research as well as to the reader who desires to find a well-conceived and thorough discussion of an important problem which cuts across many areas of psychology, physiology, and medicine.

Aero Medical Laboratory

JAMES M. VANDERPLAS

Wright Air Development Center

The Study of Language: A Survey of Linguistics and Related Disciplines in America. By JOHN B. CARROLL. Cambridge, Harvard University Press, 1953. Pp. xi, 289.

This book, we are told, is "the outcome of a report written originally at the request of the Carnegie Corporation." It is intended "to serve as a study guide" and it tries to "embrace the totality of present-day inquiry into various phases of human communication" (pp. v, vi). It is provided with an index, notes, and a list of more than 500 references, about one-tenth of them to non-English books and articles.

There are nine chapters. The first is an introduction in which the scope of the field is described. The second, which deals with the history and methodology of the science of linguistics, includes a definition of language from which we may learn something of the author's cautious attitude towards his inquiry. "A language,"

he writes, "*is a structured system of arbitrary vocal sounds and sequences of sounds which is used, or can be used, in interpersonal communication by an aggregation of human beings, and which rather exhaustively catalogs the things, events, and processes in the human environment.* The sounds and sequences of sounds are arbitrary only in the sense that they have no inherent or necessary connections with the things to which they are said to 'refer' or to the situations or contexts in which they are used—these connections having been established only by specific processes of learning" (p. 10). This definition does not refer to the origin and usage of words. Indeed, the author tells us that the scientific linguist "limits his field of inquiry by refusing to be concerned, except in a rather indirect way, with the content of the communication, that is, 'what is being talked about'" (p. 12).

In the third chapter, on linguistics and psychology, the author takes a 'neobehavioristic' position which enables him, he thinks, to reduce 'meanings' to 'learning' and 'learning' to 'reinforced associations.' What one misses is the genesis of language, either the general art of language or the acquisition of language, especially by children. There is less than satisfactory treatment of 'abstraction' and the "subjective implicit behavior" which accounts for 'thought,' 'idea,' and 'concept.' "I take the initial position," writes Carroll, "that subjective events can be regarded as behavioral, that they play an important rôle in many behavior sequences, and that appeal to them in a psychological context bears no necessary trace of philosophical dualism." "We can regard the problem of consciousness as a spurious one," he continues, "for if we think of subjective behavior as a series of responses, the notion of consciousness need not enter our discussions" (p. 72). Even so, Carroll's "guess that the selection of linguistically coded hypotheses is a function of two factors: (a) the stock of verbal responses which the subject can call up to solve the problem, and (b) the relative strengths of these responses in relation to the stimulus values of the material presented to the subject" (pp. 104-5) leaves one wondering just how relevance derives from responsiveness.

The fourth chapter considers linguistics and the social sciences, especially anthropology and sociology. It includes a discussion of an international auxiliary language with reference, chiefly, to Basic English and *Interlingua*. The fifth chapter, on linguistics and philosophy, is brief, but serves to include short reviews of the studies of C. S. Peirce and C. W. Morris. The sixth chapter, on language and education, is, like the second, a long one in which current methods and practices in the teaching of language are considered, including the foreign-language teaching-programs sponsored in the armed forces during the recent war and their effect upon present collegiate programs of study. The attitude towards these recent methods is friendly but cautious. In discussing them the remark is made that "it is also conceivable that the efficiency of a teaching method varies with the type of student. In the last analysis, we are fundamentally ignorant of the psychology of language learning" (p. 187). One might have hoped that a "study guide" based on so wide-reaching a report might refer to some specific investigations calculated to overcome this ignorance. Favorable attention is given to the views of C. K. Ogden and I. A. Richards while those of Korzybski are damned with faint praise. "Upon analysis," we are told, "the doctrines of general semantics turn out to be nothing more or less than a sprightly though somewhat muddled (*vic*) interpretation of facts and principles taught in various sciences" (p. 167). The seventh chapter, on com-

munication engineering and the study of speech, describes some of the electronic devices for recording and analyzing language. The eighth chapter, on organizations, personnel, and publications, is a factual listing of agencies.

Since the original report which furnished the basis of this book was addressed to the Carnegie Corporation, it is no surprise to find that the discussion of the future of language studies in the final chapter should be devoted to a plea for more support for the training of, and investigation by, scientific and communication "generalists" competent to handle the divers aspects of language. This plea is summarized in a list of five "priorities," the first being to collect little known languages, on the principle, perhaps, that the more data one has, the more investigations can be made. The other four priorities refer to cash—for post-doctoral scholarships and fellowships, for "interdisciplinary seminars," for "grants-in-aid," and for the support of specific researches. Are we then to suppose that more funds and more persons to spend them are all that is needed to produce a science of language? One has a feeling that something is missing from this Survey of Linguistics and Related Disciplines in America. The author has referred to almost everything in language save the fine art of literature which, many would say, is no science. Yet there it is, and remains, not the least important among "the various phases of human communication." Can it be accounted for as a reinforced association of words?

Cornell University

R. M. OGDEN

The Deprived and the Privileged. By B. M. SPINLEY. London, Routledge and Kegan Paul, 1953. Pp. vii, 208.

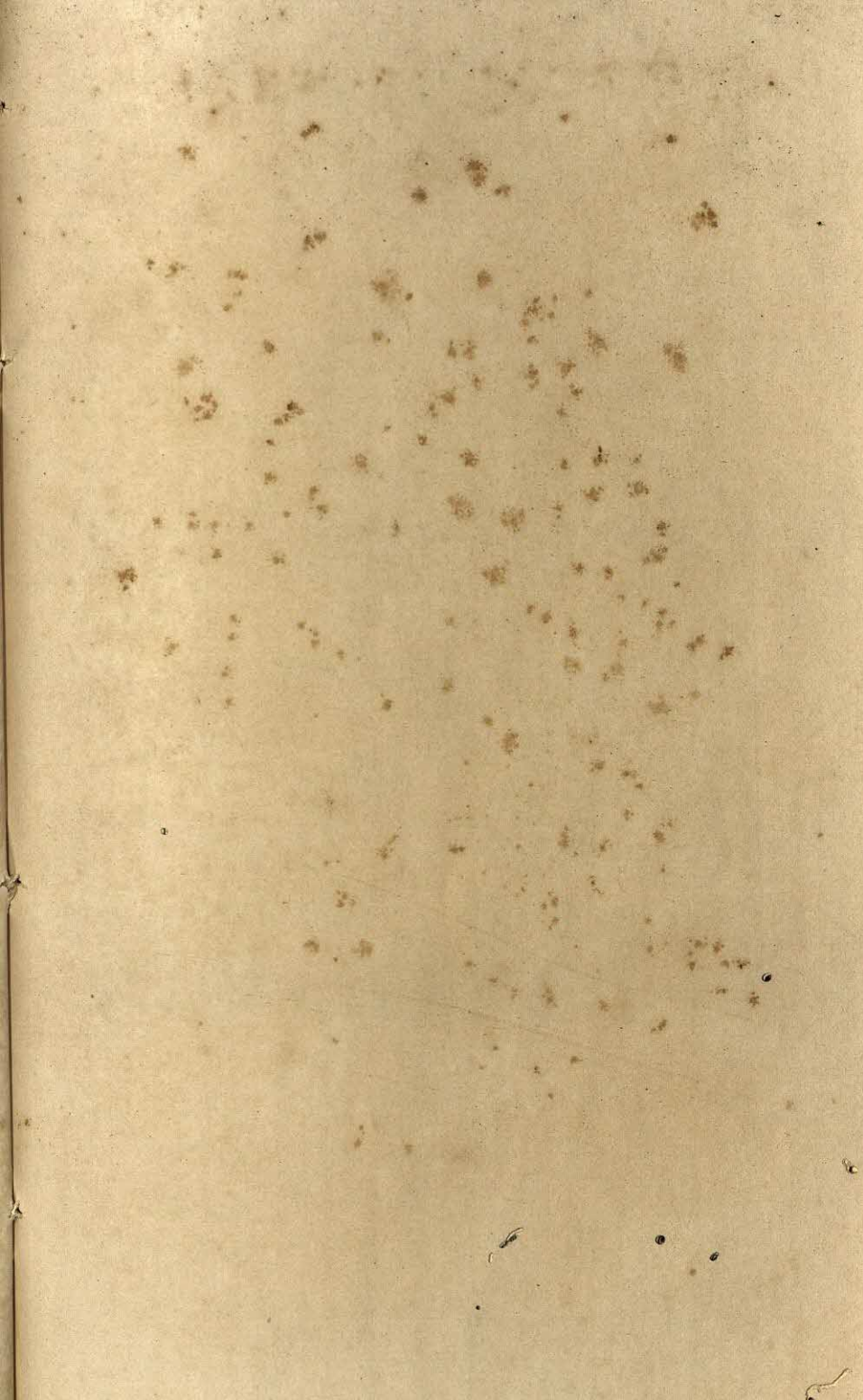
Dr. Spinley's research is an examination of two hypotheses—(1) that sub-cultures exist, and (2) that there is a basic personality type. She assumes that "differences in sub-cultural personality structures, as well as the similarities, may be associated with experiences in infancy, childhood, and possibly even the late period of adolescence." There is an historical introduction to the concepts of basic personality type and of sub-cultures, a detailed description of research procedures, a presentation of each group of results, conclusions, and appendices containing the details of the developmental studies and Rorschach analyses.

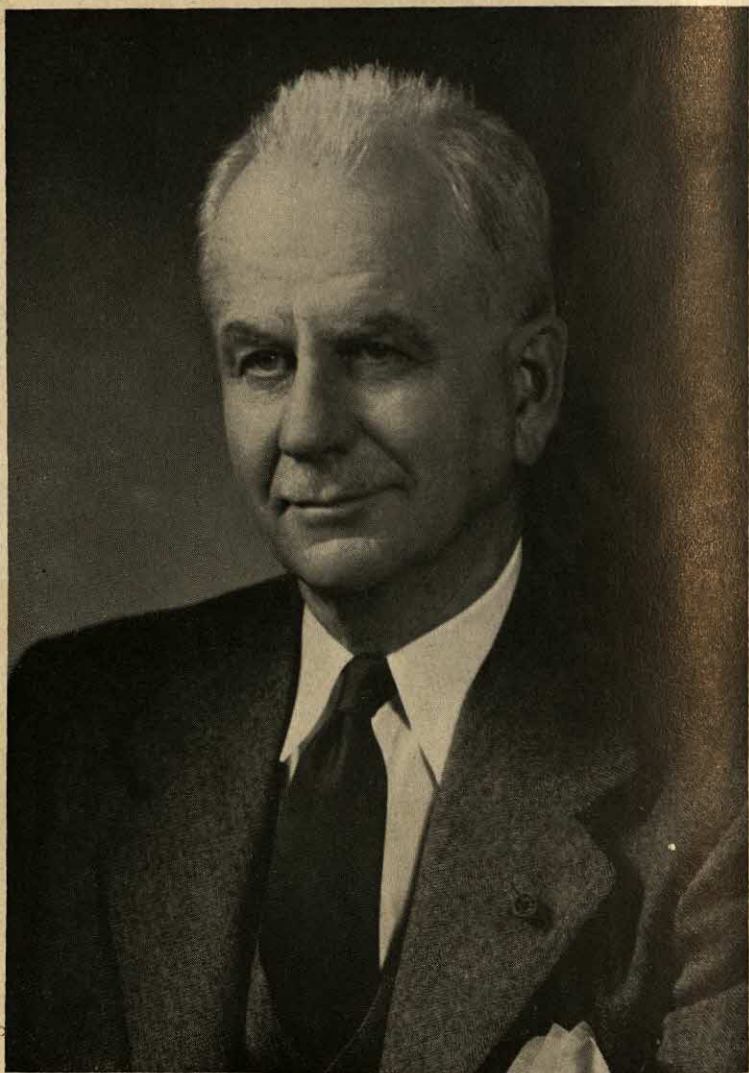
Dr. Spinley compared the developmental history of children raised in a slum area of London with that of tuition-paying children in "public schools of high repute." Data for the slum group were collected through participant observation of children in a community settlement house, through structured interviews with residents and social workers in the area, and through analysis of the Rorschach test administered to a group of school children in the area. Data for the Public School group were obtained from life histories written by former students of these public schools, through faculty interviews and study of the curriculum of "one of the large institutions which train nursery nurses," and through Rorschach tests administered to a group of Public School children matched in age and sex with the slum-area group. She concludes that her data support her basic hypotheses and that there are basic personality differences corresponding to different sub-cultural classifications.

Although the study falls short of the precision and control associated with laboratory experiments, the laboratory scientist must admire Dr. Spinley's reasonably successful attempts at control, and her caution in interpreting the data.

Knox Collège

ROBERT S. HARPER





Walter Hunter

(See page 732)

THE AMERICAN JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXVII

DECEMBER, 1954

No. 4

AN EDITORIAL ANNOUNCEMENT

Once again the editing of the JOURNAL is to become a coöperative affair. Beginning with the next number, the editorial duties and responsibilities, which I have carried alone during the past few years, will be shared with M. E. Bitterman, Visiting Professor of Psychology, University of California at Berkeley, and Edwin B. Newman, Professor of Psychology, Harvard University. Twice before my tour of duty the editorial responsibilities of the JOURNAL have been carried for short periods by individuals and twice, for longer periods, by editorial boards.

G. Stanley Hall, the founder of the JOURNAL, carried the responsibility alone through the first six volumes. After this period, and for the next twenty-five years, the editorship was a joint affair. E. C. Sanford and E. B. Titchener became co-editors in 1895 with Volume 7; and J. W. Baird, Professor of Experimental Psychology, Clark University, in 1912 with Volume 23. Baird served until his death in 1919; Sanford and Titchener until the JOURNAL was transferred to Cornell University and Titchener became its sole editor.

Titchener carried the editorial responsibilities for five years, from 1921 to 1925. He was succeeded in 1926 by M. F. Washburn, Madison Bentley, E. G. Boring, and myself—all of whom had previously served on the JOURNAL's Coöperating Board. Washburn's co-editorship continued until her death in 1939; Boring's until 1946; and Bentley's, until 1951. The vacancies caused by these successive reductions were not filled. The editorial work was merely distributed among those who remained. Bentley's retirement, consequently, left me with the full editorial responsibility of the JOURNAL.

With the generous assistance of the members of the Coöperating Board, I was able to carry on but it has become increasingly evident to me that the editing of the JOURNAL is not a spare-time job for one person. It

cannot long be carried by a single individual who has other duties to perform and other responsibilities to meet. To carry a portion of the work in one's spare time is a possibility; to carry it all is a burden that no one, as the history of the JOURNAL shows, has been able long to endure. I have no longer any doubt about the fact that the JOURNAL is served best and psychology benefits most when the editorial work is spread among several persons.

Professors Bitterman and Newman, who now form with me the third Editorial Board, were drawn, as were the members of the second Board, from the list of Coöperating Editors. They consequently bring to their new assignments a full understanding and appreciation of their work. While manuscripts and communications may be sent to any member of the new Board, promptest acknowledgments will be received from the editor living nearest to the contributor.

KARL M. DALLENBACH

THE NATURE AND HISTORY OF EXPERIMENTAL CONTROL

By EDWIN G. BORING, Harvard University

Solomon has discussed the history of the concept of control in experimental work with especial regard to the use of control groups in the design of experiments,¹ finding no instance of the employment of a control group before the study of transfer by Thorndike and Woodworth in 1901² and no extensive use of control groups in experiments on transfer until Winch's study in 1908.³ It is easily shown that the concept of control is basic to all experimental design and is, indeed, inherent in the essential relational nature of a fact. The purpose of this note is to analyze the concept more fully and to say something more about the history of both the concept and the word.

The word *control* has three meanings: (1) a *check*, in the sense of a verification but thus also in the sense of a restraint, since verification restrains; (2) a *restraint*, in the sense of a checking and thus also in the sense of maintaining constancy; and (3) a *guide* or *directing*, in the sense of producing a precisely predetermined change, a constant and thus a restrained change. The word *check* itself has the first two meanings, though not the last, and the original meaning of *control* was *check*, for the word was *counter-roll* (*contre-rolle*), a duplicate register or account made to verify an official or first-made account and thus a check by a later roll upon the earlier. (Hence *controller*, which is misspelled *comptroller* because it has been thought of as meaning an accountant instead of a checker.) So the thought of correctness or conformity achieved by restraint runs all the way through the history of the word, even though what the psychologist hears nowadays about 'controlling behavior' suggests the promotion of action more than its restriction.⁴

* Accepted for publication September 10, 1953.

¹ R. L. Solomon, An extension of control group design, *Psychol. Bull.*, 46, 1949, 137-150.

² E. L. Thorndike and R. S. Woodworth, The influence of improvement in one mental function upon the efficiency of other functions, *Psychol. Rev.*, 8, 1901, 247-261, 384-395, 553-564, esp. the eight-line paragraph on p. 558 where the use of a control group is described.

³ W. H. Winch, The transfer of improvement of memory in school-children, *Brit. J. Psychol.*, 2, 1908, 284-293. The article antedates the development of the statistical method for dealing with group differences.

⁴ Cf. B. F. Skinner, *Science and Human Behavior*, 1953, where the conception of control seems to me to be more positive than negative in the chapters entitled: The

The term *control* in the sense of a check or test observation or experiment came into scientific parlance in the latter half of the nineteenth century, as we shall see presently. By 1893 we find the *New English Dictionary* defining *control* as "a standard of comparison used to check the inferences deduced from an experiment by application of the Method of Difference," which is the name of John Stuart Mill's second method of experimental inquiry. The *New English* goes on to define *control-experiment* as "a test experiment with this end in view." So *control-experiment* was definitely in the language before *control groups* had been thought of, and this concept takes us back to Mill's four methods of experimental inquiry in his *Logic* of 1843.⁵

Mill's first method is the *Method of Agreement*: if *A* is always followed by *a*, then *A* is presumably the cause of *a*. Mere agreement does not, however, furnish rigorous proof, although you may be limited to it when you lack the voluntary variation of events—the independent experimental variable—and are reduced to description only. For this reason the establishment of causal relations in biography, history, geology, paleontology, and even astronomy is less sure than in experimental science. Mill remarked that mere agreement would indicate that night is the cause of day, and day the cause of night, since the sequence is universal, and he noted that we can be more certain that agreement indicates cause when the antecedent term in the conjunction of events can be established at will without dependence upon other events. Mill was right in mistrusting the Method of Agreement, since the concurrence of *A* and *a* in sequence may mean only that both are effects of the same sufficient cause, and, since if that other cause is sufficient but not necessary, it takes the Method of Difference to show that *A* and *a* are not necessary concomitants. It is for this reason that Mill suggested that the Method of Agreement is strengthened if *A* can be varied "at will," that is to say, if *A* is a freely independent variable. Such a caveat, however, actually constitutes an extension of the Method of Agreement to include the Method of Difference (when variation of *A* includes its elimination) or the Method of Concomitant Variation (when *A* is merely changed in degree). The inference of causation is never safe when based upon agreement alone.

The *Method of Difference* is Mill's second method: if *A* is always followed by *a*, and *not-A* is always followed by *not-a*, then *A* is certainly the cause of *a*. This is equivalent to adding the control observation: if *not-A*, then *not-a*. Mill used the word *control* once: "It thus appears that in the study of various kinds of phenomena which we can, by our voluntary agency, modify or control, we can in general satisfy the requisitions of the Method of Difference; but that by the spontaneous operations of nature those requisitions are seldom fulfilled." This use of *control* is, however, in

controlling environment; Self-control; Personal control; Group control; Economic control; Culture and control; and The problem of control. Certainly Solomon's *control group* has little relation to Skinner's *group control*. Even in discussion of experimental design the meaning of *control* is apt to vary. For instance, C. E. Buxton, in T. G. Andrews, *Methods of Psychology*, 1948, 73 f., in the course of two pages uses the word thrice as meaning maintenance of constancy of conditions and once in the sense of a test-observation.

⁵ J. S. Mill, *A System of Logic, Ratiocinative and Inductive*, 1843, Bk. III, chap. 8.

the sense of direction or guidance, whereas the Method of Difference provides *control* in the sense of a verifying check, although Mill did not make that use of the word. He recognized, however, the fundamental relation of his first two methods, speaking of the *Joint Method of Agreement and Difference*, which is essentially the modern scientific procedure for treating contingencies when continuities are not observed.

(The third method, the *Method of Residues*, need not concern us. If *ABC* is known to be the cause of *abc*, and *BC* the cause of *bc*, then *A* must be the cause of *a*, even though *A* can not be produced without *BC* nor thus *a* without *bc*.)

Mill's fourth procedure is the *Method of Concomitant Variations*. Nowadays we think of such observation as basic to all experiments and thus of Agreement and Difference as special cases of Concomitant Variation. Concomitant variation exists when there is a series of differences, and in any pair of concomitances one concomitance furnishes a comparison or control for the other. So we could get along with this method alone, if it were broadly enough conceived, except for the historical fact that the concept of control actually grew out of the consideration of the Method of Difference, which Mill's prestige established as independently important.

Mill is usually taken as the authority on this matter. Thought about these principles is, however, historically continuous. A century earlier Hume had a similar, less specific discussion, which laid down rules equivalent to the Method of Difference and to the Joint Method of Agreement and Difference.⁶ Still another century earlier Francis Bacon discussed the collection of data by the finding of instances that agree, and of negative instances and cases to furnish comparison.⁷ In these authors the anticipation of Mill is quite clear if one but remembers what a century or two can do to both thought and its expression.

At this point it is important to remark that every statement of fact expresses some kind of a difference. Even such description as is not experimental, being specific and thus discriminative, differentiates what is from what is not. Jevons offered as the formula for a fact: "Where *A* is, *X* is; and where *A* is not, *X* is not;"⁸ and that, of course, is also a statement of agreement and difference. Jevons remarked: "Every correct and conclusive experiment necessarily consists in the comparison of results between two different combinations of circumstances."⁹ It is plain then that in scientific description we are not going to get away from the concept of control, although the idea appears with different degrees of specificity and formality in statements of experimental design. If you have an observed datum, there is always some point or frame of reference in respect of which the datum makes sense, and it might aid clarity of thought if one were to think always of a datum set over against a relatum. (If *A*, then *a*) would be a datum, and (if *not-A*, then *not-a*) would be the relatum and the control observation, which might indeed depend upon a control-experiment to establish it. Again and again when agreement seems to yield certainty, it is because the control-observation is implied or even included in the experiment.

Now let us examine some instances of the use of the concept of control and of the word *control* in the scientific literature of the last one hundred years.

⁶ David Hume, *A Treatise on Human Nature*, 1739, Bk. I, pt. iii, sect. 15.

⁷ Francis Bacon, *Novem Organum*, 1620, Bk. II, Aph. 11-13, which deal respectively with the three principles.

⁸ W. S. Jevons, *Principles of Science*, 1st ed., 1874, II, 44; or 2nd ed., 1883, 433.

⁹ Jevons, *op. cit.*, 1874, II, 43; or 1883, 433.

Control as restraint or guidance. The meaning of *control* as restraint or guidance is the common, though later, meaning of the term, and in science it applies to keeping experimental conditions constant and also to altering the independent variable in accordance with precise known pre-determination. We do not need to hunt out instances of the scientific aspiration to keep conditions constant, to maintain controlled conditions or to vary a parameter under controlled conditions. In these contexts control, constancy and precision are of the essence of experimental science. Thus Titchener, in describing the nature of experiment, emphasized the necessity for repetition, isolation, and variation in the interests of good observation, constancy of conditions and exact variation, although he did not use the word *control*.¹⁰ In similar sense Murphy and Murphy have spoken of the "relatively uncontrolled observational and biographical studies" of children and the later "new experimental and highly controlled observation studies worked out" after 1915.¹¹ We can also go back to Fechner who sought constancy of conditions by control of experimental procedures and the treatment of data: "The arrangement of experimental conditions," he wrote, "the recording of observed values, the enumeration of errors or of right and wrong cases, as well as all the calculations based upon them, must be so arranged and controlled (*controliren*) by repetition and otherwise that, as far as possible, errors are avoided by the multiplication of data, by calculation when error is otherwise unavoidable, and by observing an immutable integrity in the recording and conversion of data."¹²

Control of the experimental independent variable, moreover, implies guidance as well as the maintenance of constancy, but the same admonitions and aspirations apply to this kind of control as to the control of conditions fixed throughout a particular experiment.

Control enters into psychic research with two different meanings. Holding hands and touching feet with each of your neighbors in the spiritistic circle may be thought to promote psychic continuity, but the practice also acts as a control against fraud. There is also, however, the *spirit control*, who or which guides and directs the

¹⁰ E. B. Titchener, *A Text-Book of Psychology*, 1910, 20; *A Beginner's Psychology*, 1915, 22-25. The only use of the word *control* that I can find in these books is in *Text-Book*, 22, where Titchener remarks that "experimental control is still possible" even when introspection interferes with consciousness.

¹¹ Gardner Murphy and L. B. Murphy, *Experimental Social Psychology*, 1931, 201 f., with examples of systematic arrangements for the recording of exact behavior and the maintenance of constancy of conditions on pp. 214-227.

¹² G. T. Fechner, *Elemente der Psychophysik*, I, 1860, 85; and there is another example of the use of the word farther down on the same page: "die Wiederholung oder sonstige Controle an sich langweiliger Operationen."

medium. This use of the term seems to have come in late, for Podmore makes no mention of it in 1902 in his excellent and thorough two-volume history of spiritualism.¹³ Nor did the *New English Dictionary* give this meaning in 1893, although the *Standard Dictionary* by 1913 was saying for one meaning of *control*: "The intelligence (whatever its nature) which regulates the communication of messages through a medium or psychic."

Control observations, series, and experiments. We come back to the special use of the word *control* in its original meaning as a check or standard of comparison, a relatum. If separate observations have controls interspersed (e.g. single points introduced among the double points which are used in determining the two-point limen upon the skin), then we have what may be called *control observations*. When the controls are organized into series, then we have a *control series* with which the *experimental series* is compared, as in various memory experiments. If the dissociation of the terms for comparison is greater and the organization of each more elaborate, then we begin to speak of the *principal experiment* and the *control experiment*. The term *control experiment*, as a standard of comparison, has now got into most of the dictionaries. Its synonym is *test experiment*. Logically there is, however, no difference between one kind of relatum and another except in degree of organization and formal independence. The basic conception, of course, appears early in experimentation, whereas the use of the word *control* comes later.

The concept of control is pretty old and was quite obvious once the Renaissance had turned men's thought from theological fiat to experiment as the means for penetrating into nature's secrets. Here is a story that makes the whole matter clear.

In 1648 the Torricellian vacuum was known to physics in general and to Pascal in particular. This is the vacuum formed at the upper closed end of a tube which has first been filled with mercury and then inverted with its lower open end in a dish of mercury. The column of mercury falls in the tube until it is about 30 in. high and remains there, leaving a vacuum above it. Pascal was of the opinion that the column is supported by the weight of the air that presses upon the mercury in the dish (he was right; the Torricellian tube is a barometer) and that the column should be shorter at higher altitudes where the weight of the atmosphere would be less. So he asked his brother-in-law, Perier, who was at Clermont, to perform for him the obvious experiment at the Puy-de-Dôme, a mountain in the neighborhood about 3000 ft. ("500 fathoms") high as measured from the Convent at the bottom to the mountain's top. On Saturday, September 19th, 1648, Perier, with three friends of the Clermont clergy and three laymen, two Torricellian tubes, two dishes and plenty of mercury, set out for the Puy-de-Dôme. At the foot they stopped at the Convent, set up both tubes, found the height of the column in each to be 26 old French inches plus $3\frac{1}{2}$ Paris lines (28.04 modern inches), left one tube set up at the

¹³ Frank Podmore, *Modern Spiritualism, History and Criticism*, 2 vols., 1902.

Convent with Father Chastin to watch it so as to see whether it changed during the day, disassembled the other tube and carried it to the top of the mountain, 3000 ft. above the Convent and 4800 ft. above sea-level. There they set it up again and found to their excited pleasure that the height of the mercury column was only 23 French inches and 2 Paris lines (24.71 in.), much less than it was down below just as Pascal had hoped it would be. To make sure they took measurements in five places at the top, on one side and the other of the mountain top, inside a shelter and outside, but the column heights were all the same. Then they came down, stopping on the way to take a measurement at an intermediate altitude, where the mercury column proved to be of intermediate height (26.65 in.). Back at the Convent, Father Chastin said that the other tube had not varied during the day, and then, setting up their second tube, the climbers found that it too again measured 26 in. $3\frac{1}{2}$ lines. These are reasonable determinations for these altitudes, showing about the usual one inch of change in the mercury column for every 1000 ft. of change in altitude.¹⁴

In this experiment there was no elaborate design, and it took place 195 years too soon for the experimenters to have read John Stuart Mill's *Logic*, but the principle of control and of the Method of Difference is there. How important it was for them to have left a barometer at the base of the Puy-de-Dôme to make sure that changes in the tube that they carried up the mountain were due to elevation and not to general atmospheric changes or to other unknown circumstances! How wise of the party at the top to have made the measurement under as many different conditions as they could think of with altitude constant! How intelligent of them to take a reading on the way down and thus to turn the Method of Difference into the Method of Concomitant Variation!

When Jevons was writing his section on "Blind or Test Experiments," he missed this paradigm from Pascal (1648) but gave one from Faraday (1848) and another from Tyndall (1863).¹⁵ Neither of these examples is in any formal sense a blind experiment or, as we should say today, a control experiment. They are simply examples of the basic principle that a fact is a relationship and that you are not ready to make a scientific statement until you have a comparison to present. Thus Faraday, in the research cited by Jevons as including blind experiments, was really trying out what he called magnecrystalline action on a variety of different substances, finding that certain rules of magnetic orientation hold for crystals of bismuth and similar metals and that there is no magnetic action of this kind at all for the crystals of certain other metals, of which lead is one example.¹⁶ Faraday was engaged in what

¹⁴ Blaise Pascal, *The Physical Treatises of Pascal: the Equilibrium of Liquids and the Weight of the Mass of the Air*, trans. 1937, 103-108, which is mostly Perier's letter to Pascal about the experiment. For short accounts, see I. B. Cohen, *Science, Servant of Man*, 1948, 71 f.; J. B. Conant, *Science and Common Sense*, 1951, 72-74. Cohen says that Perier noted the continuous drop in the height of the column as he came down; if three points make a continuity, he did. Conant says that Pascal was at the foot of the mountain watching the control barometer, but that was Father Chastin. Pascal was waiting in Paris to hear what happened.

¹⁵ Jevons, *op. cit.* (note 8), 1 ed., 1874, II, 43-45, is the section called "Blind or test experiments," in the chapter on "Experiment"; but the mention of Faraday's and Tyndall's experiments does not appear until 2 ed., 1883, 433 f.

¹⁶ Michael Faraday, *Experimental Researches in Electricity*, 1855, III, Series xxii, 83-136. The experiments on magnecrystalline action were performed in 1848.

might be called the experimental taxonomy of metallic behavior, and, now that his results are available to us, we can separate them into data and relata, into experiments and controls, though Faraday did not see his problem thus. Tyndall, in the place cited by Jevons, was describing an experiment in which 19 lb. of water were brought from room temperature to a boil in $2\frac{1}{2}$ hr. by the constant rotation of a steel cylinder fitting tightly inside another steel cylinder and kept by a horse in rotation while immersed in the water—a demonstration that mechanical power can be turned into heat by friction.¹⁷ Presumably Jevons believed that the control here was the initial condition of the water at room temperature, but actually this experiment was an example of the Method of Concomitant Variation, for the temperature of the water was taken from time to time and found steadily to rise as equine energy was continuously converted into heat.

In 1870 a new kind of control observation appeared in psychological experimentation, the *Vexirversuch* that Vierordt—and later Riecker in 1874—used in their determinations of the cutaneous two-point threshold.

If you keep putting two points simultaneously down upon the skin with varying separations between them and keep track of the number of times each separation is perceived as a single impression and the number of times it is perceived double, then, if your chosen separations are proper, you can compute the separation at which a *two* would be felt as often as a *one*. That is the threshold. The observer, well trained to give phenomenal reports, can participate in this experiment quite successfully, but the naïve observer is apt to make what Titchener called the stimulus-error, to report *two* often or always because he knows or guesses that the stimulator often or always puts two points to the skin. So Vierordt and then Riecker introduced single stimuli as checks upon the observer, calling them *Vexirversuche* or puzzle trials—actually control observations. Riecker worked with the method of constant stimuli which was still called then, after Fechner's original title, "the method of right and wrong cases." It is 'right' to call two points *two* and one point *one*, and 'wrong' to call two points *one* and one point *two*. Actually the observers often did call one point *two* for there is a special physiological disposition for getting a two-fold impression out of a unitary stimulus, and these errors with the *Vexirversuche* were called *Vexirfehler* (paradoxical errors).¹⁸

Later McDougall formally introduced the requirement that there should be as many single stimuli as double and that the threshold should be 80% right for both

¹⁷ John Tyndall, *Heat Considered as a Mode of Motion*, 1863, 21.

¹⁸ Karl von Vierordt, Die Abhängigkeit der Ausbildung des Raumsinnes der Haut von Beweglichkeit der Körperteile, *Zsch. Biol.*, 6, 1870, 53-72; Adolph Riecker, Versuche über den Raumsinn der Kopfhaut, *ibid.*, 10, 1874, 177-201. On this experiment, the *Vexirversuche* and the *Vexirfehler*, see Victor Henri, *Ueber die Raumwahrnehmungen des Tastsinnes*, 1898, 58-89, esp. 61-66. On the nature of *Vexirfehler*, which are no more "errors" than is any compulsive illusion, see Margaret Kincaid, An analysis of the psychometric function for the two-point limen with respect to the paradoxical error, this JOURNAL, 29, 1918, 227-232; Harry Helson and R. M. Burgert, Prediction and control of judgments from actual single-point stimulation, *ibid.*, 48, 1936, 609-616.

kinds of stimuli taken together.¹⁹ That was good functionalism at the time and is nowadays good modern behaviorism: you see how adequate the subject is to his environment (stimulus-world) when without further knowledge he relies for information upon the sensitivity of his receptor apparatus, but Titchener still thought that a trained observer should be able to judge sensory impression accurately in its own right, whatever he suspected of the stimulus.²⁰ McDougall stuck by his guns even for educated subjects, and Henry Head and his associates used McDougall's method in the experiments which set up the distinction between protopathic and epicritic sensibility.²¹ It is only recently that the *Vexirversuche* have come to be regarded as control observations, and indeed they are somewhat special for they function as straight checks or controls for objective behavior, whereas they may also become directing guides in an introspective psychology where a trained observer needs to be checked as to whether he is reporting only on sensory impression or whether, committing the stimulus-error, he is trying to infer or guess the true nature of the stimulus-object.²²

It was along in the 1870s that the word *control* began to be used in the sense of a check or a standard of comparison in respect of which a difference is expected to lie.

In 1874, in the first edition of his famous *Physiologische Psychologie*, Wundt said that the fall-apparatus, used then to calibrate the Hipp chronoscope, is a *control*. Later the fall-apparatus changed its form and became and was called a *control hammer*. Proper procedure is to run a series in which the times of the successive falls of the control hammer are measured by the Hipp chronoscope, and that series, which Wundt actually called a *Kontrollversuch*, gives you the variable error of the chronoscope and, if you have measured the true time of fall on the chronograph, the constant error too.²³

Thus Wundt was ready enough to control a piece of apparatus with an objective check, but the only control he admitted as valid for the human observer was rigorous training in psychological observation. For psychological experimentation he registered objection to *Kontrollversuche*, *Prüfungsversuche*, *Probeversuche* and *Vexirversuche*.²⁴

¹⁹ William McDougall, *Cutaneous sensation, Reports of the Cambridge Anthropological Expedition to Torres Straits*, 1903, II, pt. ii, 189-195.

²⁰ Titchener, On ethnological tests of sensation and perception, with special reference to tests of color vision and tactile discrimination in the reports of the Cambridge Anthropological Expedition to Torres Straits, *Proc. Amer. Philos. Soc.*, 55, 1916, 204-236, esp. 206-215.

²¹ Henry Head, *Studies in Neurology*, 1920, I, 26-29. The original experiment was Head, W. H. R. Rivers and James Sherren, The afferent nervous system from a new aspect, *Brain*, 28, 1905, 99-115 (reprinted in *op. cit.*, 1920, 55-65); Head and Rivers, A human experiment in nerve division, *Brain*, 31, 1908, 323-450 (reprinted in *op. cit.*, 1920, 225-329).

²² Cf. E. G. Boring, The stimulus-error, this JOURNAL, 32, 1921, 449-471, which, paradoxically, interprets Titchener's conception of the psychological point of view functionally.

²³ Wilhelm Wundt, *Grundzüge der physiologischen Psychologie*, 1st ed., 1874, 772; 6th ed., III, 1911, 367. See also Titchener, *Experimental Psychology*, II, 1905, pt. ii, 340-344.

²⁴ Wundt, *op. cit.*, 6th ed., III, 1911, 399.

Also in the 1870s this meaning of the word *control* entered biology by way of Darwin. In 1875 he reported on *Utricularia*, plants that float on foul water and that carry tiny bladders with valves that trap the insects and organic matter that get within the orifices of the bladders. It is a study in the taxonomy of plant behavior. After observing the reaction of the bladders to their normal food, "four bladders were tried as a control experiment," Darwin wrote. He fed them gum arabic and also sugar, neither of which produced the normal feeding reaction, whereas nitrate of ammonia did.²⁵ Later Darwin is found describing the movement of plants in response to irritants. He performed experiments with the radicles of peas, attaching to them as irritants bits of cardboard "which served as standards of comparison or controls."²⁶ In 1890 there was an article in *Nature* on the immunization of mice against tetanus. Immunized mice were inoculated with tetanus and failed to show its symptoms, but "control mice [not immunized] died within 48 hours."²⁷ When a word can be used thus casually, it may be assumed to have become established in the language—in this case in the vocabulary of biology.

As experimental psychology developed, the design of its experiments became more rigorous and elaborate. Speaking approximately, we may say that formal design first developed in psychophysics, then in the reaction experiments, then in the experiments on memory.

In 1900 we find Müller and Pilzecker in their classical study of right associates using comparison series (*Vergleichsreihen*) as controls (but not so named) for the principal series (*Hauptreihen*). They also introduced fore-series (*Vorreihen*) and after-series (*Nachreihen*) and thus in a way they anticipated some of the design that Thorndike and Woodworth employed shortly after for their study of transfer. You even find in Müller and Pilzecker *Hauptvorreihen*, *Vergleichsvorreihen*, *Hauptnachreihen* and *Vergleichsnachreihen*, but the *Vergleichsreihen* are control series used on the same observers. The control groups come later.²⁸

There can be no doubt that use of control observation, either implicit or explicit, is essential in sound experimental work. There has to be a relatum to give the datum significance. More definitely designed control series and control experiments have at the present time come into common use, even when control groups are not called for. Such controls go by a variety of names. For instance, in medical research one sometimes sees nowadays mention of the use of *placebo* (a dose that pleases the patient but has no pharmacological effect; *placere*, to please) introduced as a control in comparison observations or in control series.²⁹

²⁵ Charles Darwin, *Insectivorous Plants*, 1875, chap. xvii, esp. 413.

²⁶ Charles Darwin and Francis Darwin, *Power of Movement in Plants*, 1881, chap. iii, esp. 160-163.

²⁷ E. H. Hankin, A cure for tetanus and diphtheria, *Nature*, 43, 1890, 121-123.

²⁸ G. E. Müller and Alfons Pilzecker, Experimentelle Beiträge zur Lehre vom Gedächtniss, *Zsch. Psychol.*, Ergbd. 1, 1900.

²⁹ Stewart Wolf, Effects of suggestion and conditioning on the action of chemical

When there is no acceptable hypothesis as to the independent variable in an experiment, it may be impossible to have a control, which may become available only when a good hypothesis comes along. Thus, in the experiments on extrasensory perception (ESP), control is very difficult. Control by restraint and guidance is common enough for every effort is made to keep conditions constant and to prevent fortuitous leakage of information to the percipient, but control as a check requires a knowledge of what is to be omitted if the Method of Difference is to be used. The reading of cards in an invisible pack is supposed to have been extrasensory when the reading deviates from 'chance' by an amount that can be accepted under current statistical conventions with a high level of confidence. But what is 'chance'? Good statisticians nowadays do not accept, as they did thirty years ago, the principle of insufficient reason as meaning the a priori presumption that wholly unconstrained coincidence will give correct guesses in only 20% of the cases when there are equal numbers of five kinds of cards in the pack. In this case there needs to be a control experiment, one that would show empirically what the 'purely chance' frequencies are. Usually the parapsychologists attempt such controls, but the ideal is to have the same subject guess the same pack of cards both with and without ESP. How do you turn off ESP in a subject and turn it on again and know what it is you have done? It is not impossible that some of the heated argument about ESP would disappear if there could be a clearer understanding of what a control is and when one is being used.³⁰

Control groups and the measurement of the differences between groups. There is an advantage in using an entirely different set of subjects in the control experiment from those employed in the principal experiment.

agents in human subjects—the pharmacology of placebos, *J. Clin. Investigation*, 29, 1950, 100-109; F. K. Abbot, Marian Mack and Stewart Wolf, The action of bantnine on the stomach and duodenum of man with observations on the effects of placebos, *Gastroenterology*, 20, 1952, 249-261; J. D. Hardy, H. G. Wolff and Helen Goodell, *Pain Sensations and Reactions*, 1952, 231.

³⁰ J. G. Pratt, J. B. Rhine, B. M. Smith, C. E. Stuart and J. A. Greenwood, *Extrasensory Perception After Sixty Years*, 1940, esp. chaps. 1-3. P. 26 says: "It can be demonstrated mathematically that so long as the target deck patterns contain five of each design, the average number of correct matchings with any call pattern is invariably five." You can not, however, demonstrate any empirical event mathematically. All you can show mathematically in this case is that 1/5 of the possible patterns would give 1/5 of the matchings correct, but that is not to say that the actual patterns would average out to 1/5 correct matchings. For a given call pattern the decks could be rigged to give any desired frequency of correctness. A control experiment might show and might not show what effect ESP has on frequency of correct matching, depending on how much in each of the two situations to be compared is left uncontrolled (in the other sense of that word). On the difficulty of using the probability model to fit an empirical situation, see Boring, Statistical frequencies as dynamic equilibria, *Psychol. Rev.*, 48, 1941, 279-301.

When the same subjects are used in both experiments, you can not be sure that the one experiment, whichever is performed first, has not affected the other, and especially is that danger great in experiments upon learning where it is expected that initial events will affect later ones, since that consequence is exactly what learning is. It is no wonder then that the first use of control groups in psychological investigation came about in experiments on the transfer of training, nor that the control group is now so prevalent in the design of psychological investigations, since learning is at present so frequent a topic of research and rats, cheap and plentiful, are the chief class of subjects.

There is, however, also a disadvantage that comes with the use of the independent control group. The use of separate subjects makes sure that there is no spurious cross-effect between the principal experiment and the control, for Peter's memories do not reflect Paul's learnings when there has been no social contact between the two. Thus you are safe in respect of the parameter under investigation; but what, one may ask, about other parameters? How do you know that your control group is the equivalent of your principal group in every relevant attribute except in respect of the independent variable, for which a difference is intended and assured? You have to have equivalent groups. You may match them individual for individual in respect of what seem to be their most important determinable and presumably relevant characteristics. You may make both groups very large, hoping that the law of large numbers will change your ignorance into assurance of equivalence. You can match litter-mates in body-weight if your subjects are animals, and you can advertise for twins when your subjects are human. Still your subjects in the two groups will not be identical, and you must on that account accept a loss in confidence to set over against the gain that led you to resort to control groups in the first place. There lies here a principle of complementarity from which you can not escape.³¹

Since Solomon has already traced the early history of the introduction of the control group into experimental psychology, there is no need for us to repeat his comments, and we may rest content with recapitulating his important points.³²

³¹ T. G. Andrews, *Methods of Psychology*, 1948, 10-15, distinguishes five methods of control: (1) the control test, which is simply the Method of Difference; (2) the control group as described above; (3) matched pairs, in which equivalence of control groups is assured by matching every individual in the experimental group to one in the control group in respect of all conceivably relevant parameters except the crucial one under investigation; (4) the practice method, in which learning is made maximal before the experiment so that further improvement in the fore-test is improbable; and (5) the rotation method where the orders of tests are counterbalanced, for instance, the tests A, B and C are given for one group in the order ABC, for another in the order CAB, and for a third in the order BCA. When these orders are combined, it is hoped that practice effects due to priority will cancel out, but, of course, you can not be sure. Counterbalancing is not constancy, and the principle of complementarity applies here too. Similarly the method of practice gives up the attempt to study effects of early practice.

³² Solomon, *op. cit.* (note 1).

The earliest example of a control group mentioned in the present paper is the control mice who died of tetanus because they were not immunized, but that is not a psychological investigation.³³ Solomon makes much of the fact that Thorndike and Woodworth were the first to see the need for a control group in the study of the transfer of training, but they did not fully appreciate its importance, for what they describe is only an incidental control experiment, to which they give eight lines of description in a paper of 1400 lines.³⁴ The classical experiments on transfer without a control group are open to the criticism that you can not tell how much practice in skill *A* affects skill *B* unless you have tested *B* before the practice, and then this initial testing of *B* may itself have acted as practice on *B* to vitiate the results. You can measure this special practice effect of the initial testing on a control group, however, and then accept as significant the difference between the groups. Examples of well known investigations of transfer without control groups are Volkman's on bilateral transfer long ago in 1858,³⁵ and much later Swift's³⁶ and Ebert and Meumann's.³⁷ Dearborn criticized these papers, showing that improvement in the after-test was due largely to learning in the fore-test and not, as had been supposed, to the intervening practice on the alternative task.³⁸ The first investigation to use the full design of (fore-test→practice→after-test) in the experimental group, to be compared with (fore-test→nothing→after-test) in the control group was Winch in 1908, and the general use of control groups in these experiments begins then.³⁹ Sleight used a control group (not so named) in 1911.⁴⁰ Later Ewert's study of bilateral transfer is representative of good work,⁴¹ and Bray gives a good historical account of experimental design in bilateral transfer.⁴² Woodworth's 1938 summary of the requirements in studying transfer is well known.⁴³

As we have already remarked it is hard to establish causal relations in the study of history, for we have there only the Method of Agreement and no controls. Nevertheless historical agreements are interesting. Even if they do not show that one event is cause and another effect, they tend often to support the *Zeitgeist* conception of progress by showing that related conceptions are apt to emerge at about the same time. Thus we may note as a pertinent example of historical simultaneity that the

³³ Hankin, *op. cit.* (note 27).

³⁴ Thorndike and Woodworth, *op. cit.* (note 2).

³⁵ A. W. Volkman, Ueber den Einfluss der Uebung auf das Erkennen räumlicher Distanzen, *Ber. sächs. Akad. Wiss. Leipzig, math. phys. Cl.*, 10, 1858, 38-69.

³⁶ E. J. Swift, Studies in the psychology and physiology of learning, this JOURNAL, 14, 1903, 201-251.

³⁷ Ernst Ebert and Ernst Meumann, Ueber einige Grundfragen der Psychologie der Uebungsphänomene im Bereiche des Gedächtnisses, *Arch. ges. Psychol.*, 4, 1904, 1-232.

³⁸ W. F. Dearborn, The general effects of spaced practice on memory, *Psychol. Bull.*, 6, 1909, 44.

³⁹ W. H. Winch, The transfer of improvement of memory in school-children, *Brit. J. Psychol.*, 2, 1908, 284-293.

⁴⁰ W. G. Sleight, Memory and formal training, *ibid.*, 4, 1911, 386-451.

⁴¹ P. H. Ewert, Bilateral transfer in mirror-drawing, *Ped. Sem. (and J. Genet. Psychol.)*, 33, 1926, 235-249.

⁴² C. W. Bray, Transfer of learning, *J. exper. Psychol.*, 11, 1928, 443-467, esp. 443-446 for historical account.

⁴³ R. S. Woodworth, *Experimental Psychology*, 1938, 178-181. In the new edition, see the long chapter on transfer, Woodworth and Harold Schlosberg, *idem*, rev. ed., 1954, 733-778, esp. 735-738.

use of group controls requires the comparison of two groups and that the statistical techniques for measuring the significance of group differences were being invented or discovered just about the time group controls came into general use. The year 1908 saw an early important paper in both lines of development—Winch's experiment with group controls and "Student's" paper on the significance of differences.

The history of the measurement of differences between groups runs approximately like this. The mathematicians—LaPlace, Gauss, Bertrand, Poincaré—and the early statisticians—Quetelet and Galton—all had the formulas for the probable error or standard error of a mean and for the variability of the difference between two such means. They sometimes wrote the ratio of the difference to its variability, which was much later called the critical ratio.⁴⁴ In England in the last quarter of the nineteenth century there was great interest in the study of quantitative relations between groups, but concern centered on regression and correlation and not on differences. Galton's best known paper on regression appeared in 1886 and his paper on correlation in 1888, whereas Pearson first described product-moment correlation in 1896. Then in 1900 Pearson published what was to become a basic paper in the measurements of group differences. He called it: "On the criterion that a given system of deviations from a probable case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling."⁴⁵ Pearson gave here a chi-square table and prepared the way for what came later. His problem was really the one that R. A. Fisher took over much later: How significant is the deviation of a mean from the null hypothesis? The crucial though not the first paper was, however, the one in 1908 by the modest "Student," who only much later came to be identified as W. S. Gosset. It dealt with the probable errors of means and of the differences of means and gave the famous *t*-tables for assessing the significance of differences.⁴⁶ Pearson published again on this matter in 1911,⁴⁷ and the practical measurement of differences and the assessment of their significance was well under way. It is interesting to note that in those days—the decade of the 1910s—Pearson used six-place tables and a difference was regarded as 'certain' only if the probability of its being in the direction found was 1.000000, that is to say, somewhat greater than 0.999999 or less than one chance in a million of being wrong. It was not until the 1920s that the three-place tables came to be accepted for this purpose and that McGaughey, in 1924, could assert that it was safe to accept differences when the critical ratio is only 3 or greater.⁴⁸ Of course, nowadays we all follow R. A. Fisher who in the 1930s introduced the concepts of degrees of freedom, the null hypothesis, fiducial limits and levels of confidence.⁴⁹

⁴⁴ H. M. Walker, *Studies in the History of Statistical Method*, 1929, 55 f., 199; Victor Henri, *Le calcul des probabilités en psychologie*, *L'Année psychol.*, 2, 1895, 466-500 (passim). The statement in the text is too general, but it serves to set the stage for the late 1890s.

⁴⁵ Karl Pearson, On the criterion that a given system of deviations from a probable case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, *Philos. Mag.*, 50, 1900, 157-175.

⁴⁶ "Student" [W. S. Gosset], The probable error of a mean, *Biometrika*, 6, 1908, 1-25.

⁴⁷ Pearson, On the probability that two independent distributions of frequency are really samples from the same population, *ibid.*, 8, 1911, 250-254.

⁴⁸ J. R. McGaughey, *Fiscal Administration of City School Systems*, 1924, 9, 31.

⁴⁹ R. A. Fisher, *Design of Experiments*, 1st ed., 1935, chap. 10 et passim, and so in later eds., 1937, 1942, 1947.

That is a rough sketch of what was going on at this time in respect of the measurement of differences between groups. That the problem was in the air and that the times were right for its advancement is further indicated by the fact that Victor Henri in Paris in 1898 published a paper on the probability of differences, referring to a table of probabilities for critical ratios which he had already printed in his earlier and more fundamental paper of 1895.⁵⁰ This 1898 paper of Henri's anticipated Pearson's similar one by two years but it received almost no attention, for the main movement was going on in London and in English, and Henri's voice was little more than a murmur in remote Paris. Besides in France there was no one to push this sort of investigation, whereas London after a bit had the new Galton Laboratory. Nineteen years afterward an American psychologist did discover Henri's paper. He commented on it, corrected it, explained it, and invented the phrase "probable correctness of a difference," but his paper too was peripheral to the main stream and perhaps no one but he, his wife and the editor ever read it.⁵¹ Still these eddies at the side show how the main stream is flowing, out in the center of history. That is the way the Zeitgeist works.

Now we ask: Did the use of group controls wait upon the development of the statistical methods for measuring differences between groups? That is a hard question to answer, for, as is usual in historical questions, we have only Agreement for evidence and no control observation. Winch with his good design for group control in 1908 used none of the new statistics. He compared individuals in the groups case by case and drew his conclusions. Besides Gosset did not even publish his paper until 1908 and it took time and help from the Galton Laboratory before psychologists found out how to use it. It seems instead that the initial relation between the two developments was coördinate, not causal. Nowadays, however, when scientific assurance can be had for only a 5-% chance of being wrong and truth exists at the 1-% fiducial level, half the experimental papers can use group controls and many of them use the new statistics. Even though neither development caused the other, there has certainly been a circular positive feed-back between the two technics, one that helps to keep them both in use.

Prevalence of control, explicit and implicit in scientific activity. We

⁵⁰ Henri, Quelques applications du calcul des probabilités à la psychologie, *L'Année psychol.*, 5, 1898, 153-160. For Henri's 1895 paper, see note 44.

⁵¹ Boring, On the computation of the probable correctness of differences, this JOURNAL, 28, 1917, 454-459.

shall say little here of *control* in the sense of restraint or guidance. Of course, experimental conditions must be kept constant. Every experiment has an independent variable and preferably only one, though an experimental design may involve many experiments with different independent variables. The independent variable must be constrained to vary in a predetermined way, and all other possible variables must be kept constant. The following paragraphs, however, consider *control* only as a check and as it is related to the Method of Difference.

At first it seemed worth while to find out how much psychologists are thinking about controls in their experimental work, so I took one of the 1951 volumes of the *Journal of Experimental Psychology* (vol. 41) to assess the situation (Table I, columns 4 and 7). I was astonished to find that 52% of the 58 experimental articles in this volume used control groups, so I analysed the contents of the 1916 and 1933 volumes of the

TABLE I

FREQUENCIES OF EXPERIMENTAL ARTICLES IN WHICH CHIEF FACTS DEPEND ON USE OF CONTROLS OR OF OBSERVED DIFFERENCES IN THREE VOLUMES OF THE JOURNAL OF EXPERIMENTAL PSYCHOLOGY

	Frequencies			Percentages		
	Vol. 1 1916	Vol. 16 1933	Vol. 41 1951	Vol. 1 1916	Vol. 16 1933	Vol. 41 1951
(1) Control groups; group differences	0	6	30	0	11	52
(2) Control series; control observations; observed differences	4	18	16	17	34	27
(3) Concomitant variation	15	18	11	62	34	19
(4) Factual description	5	11	1	21	21	2
Total	24	53	58	100	100	100

same journal. No control groups were used in 1916, only 11% in 1933. I also checked on the first volume of the *American Journal of Psychology* (1887) and, to get something simple before experimental psychology existed, volume 73 of Gilbert's *Annalen der Physik* (1823), which most psychologists know in later years as Poggendorf's *Annalen*. There were no group controls in them. That is, of course, as expected.

Next I counted experimental articles which had in them explicit reference to control series or control observations or which were concerned with establishing a difference between two conditions (not individual differences). It is impossible to distinguish these two kinds of differences between compared conditions—as to whether one condition provides a second datum (case of two principal sets of observations) or merely a

relatum (case of principal observation and control observation). Line 2 of the Table shows these frequencies. When three or more conditions are compared, you come to the various experimental patterns which may be called concomitant variation, some of them quite complicated. See line 3 of the Table. About three-fourths of the experimental articles in the 1887 *American Journal* and half the experimental articles in the 1823 *Annalen* were of these two kinds: studies of a pair of concomitant differences or of more involved concomitant variations.

There are also the experimental articles which provide factual descriptions like determinations of the predominant frequencies of vowel sounds, or the spectral specification of the stimulus to white, or the examination of the sensitivity of the blind spot, or the establishment of the cues to the perception of moisture (line 4 of the Table). These descriptions constitute a kind of experimental taxonomy and are characteristic of the early stages of the development of an experimental field. Nearly half of the experimental articles in the 1823 *Annalen* are of this sort, a fifth of the articles in the 1916 and 1933 *Journal of Experimental Psychology*, but only one article out of 58 in 1951. What has been happening in experimental psychology is that psychologists, perhaps especially under the stimulus of R. A. Fisher, have been becoming more conscious of experimental design. Thus they are aware of the relata, and either they investigate differences between phenomena or they introduce experimental control observations, making the comparison of events explicit and measuring the significance of the difference between them.

In making these counts I have ignored non-experimental articles—factual discussion, theoretical discussion, polemics, and papers describing apparatus, techniques and methods. In the volumes examined these kinds of articles are more common in the earlier years and that seems proper in view of what we know about the history of sciences, although particular editorial policy also has something to do with these frequencies.

¹⁰ The prevalence of group controls in recent psychological experimentation is undoubtedly due in part to the greater availability of subjects. In Germany and America in the 1890s the students worked with each other and even with the professor as observers. Such studies remain important but are supplemented nowadays by researches which utilize large numbers of rats and school-children, subjects which, besides being inexpensive and plentiful, make the sample examined more representative of larger populations and less dependent upon possible idiosyncrasies.

SUMMARY

Control in the sense of restraint has always been used in experimentation to keep conditions constant and is thus an essential part of the experimental method.

Control in the sense of guidance is involved in causing an independent variable to vary in a specified and known manner and is thus also essential in experimentation.

Control in the sense of a check or comparison, the original meaning of the word, appears in all experimentation because a discoverable fact is a difference or a relation, and a discovered datum has significance only as it is related to a frame of reference, to a relatum.

The methodological status of control as a check or comparison is to be understood by reference to J. S. Mill's Method of Difference.

Control or the frame of reference is often only implicit, as it is in factual description.

The control observation is often only implicit, but the control series or the control experiment is generally explicit, being taken into consideration in the design of an experimental investigation.

The explicit control experiment is not new. Pascal used it in determining the weight of the air in 1648.

The experimental use of control groups is new, being found in biological research in the late nineteenth century and in psychological research since the early twentieth century. Its use at the present time, especially with animal subjects and school-children, is very common.

Although the use of the control experiment is early, the use of the word *control* in the sense of a check or comparison does not seem to appear in the scientific literature until the latter part of the nineteenth century.

TESTS OF MOTOR FUNCTIONS IN INVESTIGATIONS ON FITNESS

By JOSEF BROŽEK and HENRY L. TAYLOR, University of Minnesota

Man's 'fitness' represents an integral of many functions and its comprehensive, quantitative description is an interdisciplinary task involving somatic morphology (physical anthropology), biochemistry, physiology, and psychology.¹ The study of motor functions, shared by physiology and psychology, represents a practically important segment of investigations on fitness and the changes in fitness under the impact of a variety of biological stresses.²

In the context of nutritional, climatic, pharmacological, and other stresses, the investigator concerned with manipulative performance is faced from the outset with a difficult choice between complex, concrete performances closely tied with man's occupational activities, and the use of relatively simple, factorially 'pure,' abstract tests. Only rarely, principally in connection with military applications, is it feasible to use as a criterion the scores obtained in 'field performance,' such as firing a rifle.³ Miniature work-situations constitute an attractive and promising research-tool, deserving greater attention in the study of deterioration of fitness under stress than has been paid to them. They represent a workable compromise between the realism of industrial and military field-studies, and the controlled, systematic, componental laboratory research.⁴ One distinct advantage of miniature work-situations is that the performance can be continued over a period of hours and thus the test can serve as a measure of endurance; the usual laboratory tests of motor functions are short, rarely exceeding 5 min. In a study on the effects of illumination, using a test reproducing the

* Accepted for publication November 19, 1953. This study was supported in part by funds provided under contract AF-33-038-21914 with the USAF School of Aviation Medicine, Randolph Field, Texas. The technical assistance of Joseph C. Franklin, Harold Guetzkov, and Ersal Kindel is acknowledged.

¹ H. L. Taylor and Josef Brožek, Evaluation of fitness, *Fed. Proc.*, 3, 1944, 216-222.

² Ancel Keys, Brožek, Austin Henschel, Olaf Mickelsen, and Taylor, *The Biology of Human Starvation*, 1950, 1071-1104.

³ D. B. Tyler, Effect on marksmanship of some motion sickness preparations containing barbiturates and hyoscine, *J. Appl. Physiol.*, 1, 1949, 737-742.

⁴ Brožek and J. V. Monke, Miniature work situations as a research tool, with special reference to the study of physiological stresses, *Arch. Indust. Hyg. Occup. Med.*, 2, 1950, 63-97.

essential features of the job of inspecting objects transported on a conveyor,⁵ the performance-criteria (level of performance or work-decrement) were more sensitive to variations in the level of illumination than any of the large number of sensory or motor functions measured during or before-and-after a 2-hr. period of strenuous visual work.⁶ It is likely that in other types of stresses the experience would be similar. Psychomotor tests of endurance of the analytical type, such as tapping at maximal speed, have little or no similarity to this aspect of performance as it is known in daily tasks; the test-performance involves small muscle groups and a measurable deterioration sets in in a matter of seconds. Practically useful endurance-tests have been designed to give information on the capacity to perform strenuous physical activity. These tasks (*e.g.* running on an inclined, motor-driven treadmill) appreciably increase the total energy output and it is the cardio-respiratory rather than neuromuscular functions which constitute the limit of performance.

While the miniature work-situations have important advantages as tools of research on sensorimotor functions, their construction and application tend to be time-consuming and costly. These work-tests are likely to take a good deal of laboratory-space. When the performance involves a significant amount of motor skill, long-continued improvement due to practice (versus deterioration under stress) poses difficult problems. The results are meaningful for the concrete situation of which the test is a replica but the very specificity of the task severely limits the possibilities of generalization. All these and other conditions led us to the choice of the componental study of motor performance, after a brief exploration of some synthetic tasks.

The battery of tests described here grew out of the needs and in the physical framework of our Laboratory. We believe, nevertheless, that the information on the effects of practice, on the consistency of repeated measurements, on the inter-relationships between motor functions, and on the sensitivity of the different tests and functions to stress will be of interest and value to others faced with similar problems. The relative deterioration under a variety of biological stresses constitutes the core of this paper.

DESCRIPTION OF THE TEST-BATTERY

The standard battery included tests measuring strength (two tests), speed (three tests), and eye-hand coordination. Strength was measured by a standard hand-grip

⁵ Brožek, Ernst Simonson, and Keys, A work test for quantitative study of visual performance and fatigue, *J. Appl. Psychol.*, 31, 1947, 519-532.

⁶ Brožek, Simonson, and Keys, Changes in performance and in ocular functions resulting from strenuous visual inspection, this JOURNAL, 63, 1950, 51-66.

dynamometer of the rectangular type,⁷ and a back-pull dynamometer with a horizontal handle bar attached by a chain to the spring-scale mounted on a foot-board.⁸ The number of chain links was adjusted to S's height.

(a) *Coördination*. In the test of the speed of coördinated hand and arm movements,⁹ S held a steel pipe (30.5 cm. in length and 1.5 cm. in diameter) in one hand vertically in front of himself. With the other hand he placed a steel ball at the top of the pipe, let it drop through and caught it at the bottom, repeating the performance as rapidly as possible for 1 min. A safety net was suspended under the pipe that S could readily pick up the ball if it were not caught. A mechanical counter, with a level extending into the pipe, was activated by the falling ball. The number of times the ball passed through the pipe in 60 sec. served as the score. The performance required fast vertical movements of the forearm, combined with pronation and supination of the hand. The fingers participated in the placing and the catching of the ball. Visual cues became so unimportant after a short period of practice that the task could be performed with the eyes closed. In the standard procedure the test was performed while S walked on a motor-driven treadmill. This increased the element of coördination and made the test more difficult.

The remaining three tests were also performed on the treadmill.

(b) *Tapping*. The speed of small repetitive hand-movements was measured by a 2-plate tapping test, which was used rather than the one-plate form to prevent artifacts due to tremor. The score was the number of taps in 10 sec.

The tapping-plates and the pattern traced with a stylus were mounted on a board, located in front of S. When not in use, the board was raised out of the way. The tapping plates (4 x 5 in.) were separated by a barrier ($\frac{3}{16}$ in. thick and 1 in. high), insulated from the tapping plates. The tapping-stylus was tipped by a steel ball $\frac{1}{4}$ in. in diameter. A contact between the stylus and the plate was registered by a Cenco counter. The duration of the test was determined precisely by a Telechrone motor with a contact-arm moving over a surface containing three metal contact-plates. As the arm moved, a circuit was momentarily closed and a 'ready' signal (amber light) located on the left side of the tapping-plate was turned on. Two seconds later a 'go' signal (green light) appeared, indicating that S should begin tapping. S tapped for 14 sec. until the appearance of the 'stop' signal (red light), but the counter only recorded the taps for the middle 10 sec.

(c) *Tracing*. Tracing a grooved pattern with a stylus was used to measure eye-hand coördination. Both the number of the involuntary contacts between the stylus and the sides of the pattern, and the total duration of the contacts, recorded electrically, were used as performance scores.

The pattern to be traced (10 x 10 in.) was mounted at an angle of 30° from the horizontal. S was trained to trace the pattern in a standard time (twice in 45 sec.). The groove was $\frac{3}{8}$ in. wide and $\frac{3}{8}$ in. deep. The bottom of the pattern consisted of a sheet of stainless steel; the sides, insulated from the bottom, were made of silver strips. When the stylus, similar to the tapping stylus, touched either side a contact was registered by the counter and the timer. A buzzer indicated that S was touching the side of the groove. A separate stylus was used in the tracing

⁷ C. H. McCloy, *Tests and Measurements in Health and Physical Education*, 1942, 28-29.

⁸ McCloy, *op. cit.*, 30.

⁹ Brožek, A new group test of manual skill, *J. Gen. Psychol.*, 31, 1944, 125-128.

test as the one used in tapping was flattened in the process. This would make the pattern tracing easier and result in 'better' scores simulating a continued practice improvement. (This is an example of artifacts against which *E* in longitudinal experiments must be constantly on guard.)

(d) *Reaction-time*. The test of reaction-time involved a discriminative response to simple visual stimuli. *S* had to bend his knees and strike the appropriate key (green, left key; red, right key; white, both keys simultaneously). The stimulus-lights were presented in random order and in an irregular time-sequence. The score was the average time, in 0.01 sec., of 25 reactions.

The stimuli for the reaction-time test were provided by three lights, with the bulbs mounted in front of *S* at a distance of about 5 ft. A blind in front of the lights was adjusted in height in such a way that *S* could just see the lights when walking erect. This prevented anticipatory reactions from a stooped position. *S* reacted by depressing the proper key. The keys were mounted on each side of the treadmill and were adjustable to *S*'s height. They were made in the form of large paddles to minimize the element of precision in the execution of the movements. There was a delay of at least one-half second after the key was struck, to allow *S* to regain the upright position before the next stimulus appeared. The time for 25 reactions was recorded automatically.

Two checks for standardization of the recording instruments were built into the equipment. The timer received 40 impulses in 10 sec., each impulse lasting 0.1 sec. The total time accumulated on the timer, when the electronic relay and the clutch of the timer are in good order, must give a constant value (4 sec.). The counter received 100 impulses in 10 sec.; this rate exceeded the maximal tapping speed.

For *E*'s convenience the last three tests (tapping, pattern-tracing, reaction-time), were operated from a central control panel. The apparatus was so designed that a turn of the selector switch was all that was necessary to make all the proper connections. By pushing a single button the presentation of stimuli and recording of responses was set into motion. The recording instruments and the instruments for an automatic presentation of stimuli (in the reaction-time test) or for timing of the tests (in the tapping test) were made in duplicate, to prevent loss of valuable data due to breakdown of equipment. There were two separate control panels containing a timer and an impulse counter as well as the selector switch (in the center), a voltmeter, the start and stop push-buttons (turning on a red and green light on the panel), and a general switch connecting the apparatus to the source of alternating current. All the relays and motors were placed in two noise-proof cabinets, located below the control panels, to prevent *S* from receiving audible cues. The apparatus was as complex as its operation was simple. Simplicity of operation, together with the sturdiness of construction and reliability of recording, was among the chief aims of the design. It made it possible for *E* to observe *S* during the test and to remain in personal contact with him.

In addition to the tests described in the present paper, other procedures for the study of motor functions were developed.¹⁰ These included a toe reaction-time designed for research on deficiency of the B-complex vitamins; a test of lift-strength for the leg; a method for the analysis of voluntary eye-movements;¹¹ and

¹⁰ Keys et al., *op. cit.*, 1950, 1095.

¹¹ Brožek, Quantitative criteria of oculomotor performance and fatigue, *J. Appl. Physiol.*, 2, 1949, 247-260; Quantitative analysis of voluntary eye movements, *Methods in Medical Research*, 3, 1950, 199-207.

a test of speed of non-repetitive leg and arm movements.¹² A measure of body sway (ataxiometer), a paper-and-pencil modification of the two-plate tapping test; and a test of handwriting, evaluated as motor performance not as a 'projective' personality test, were also used in some of the experimental series.

EFFECTS OF PRACTICE

In long-term experiments, which require repeated measurements of characteristics of fitness susceptible to improvement with practice and in which the Ss frequently serve as their own controls, it is necessary to establish 'normal baselines' before the start of the experimental period proper. Information on the rate of improvement with practice and on the time necessary for reaching the plateau is very important for the planning of the experiments.

Three sets of data were obtained. The first set, Series 1, included an earlier version of the test-battery described elsewhere.¹³ The second and third set, Series 2 and 3, were obtained with the battery described in the present paper. In Series 1, three measurements were made on the strength tests and two on the other tests,

TABLE I
APPROXIMATE NUMBER OF SESSIONS NEEDED TO REACH A FAIRLY
STABLE PLATEAU OF PERFORMANCE

Test	Series		
	1	2	3
Handgrip dynamometer	1	6	4
Back-pull dynamometer	2	6	4
Tapping	5	9	(4)*
Speed of hand and arm movements (ball-pipe test)	7	15	(7)†
Reaction-time	10	14	13
Pattern-tracing, duration of contacts	14	11	18
Pattern-tracing, number of contacts	14	10	18

* Stationary paper-and-pencil test. In Series 1 and 2 the task was performed on the treadmill.

† Performed in groups of 6 Ss at a time and while standing. In Series 1 and 2, every S was measured individually on a moving treadmill.

during the practice-trials in the morning and afternoon. In Series 2 and 3 the present test-battery was given three times during the morning and afternoon sessions and the measurements were averaged. These averages are shown in Figs. 1-3 and are referred to in Table I.

¹² Brožek, Simonson, and Keys, A test of speed of leg and arm movements, *J. Appl. Physiol.*, 4, 1952, 753-760.

¹³ Brožek, Harold Guetzkow, Mickelsen, and Keys, Motor performance of normal young men maintained on restricted intakes of vitamin B complex, *J. Appl. Psychol.*, 30, 1946, 359-379.

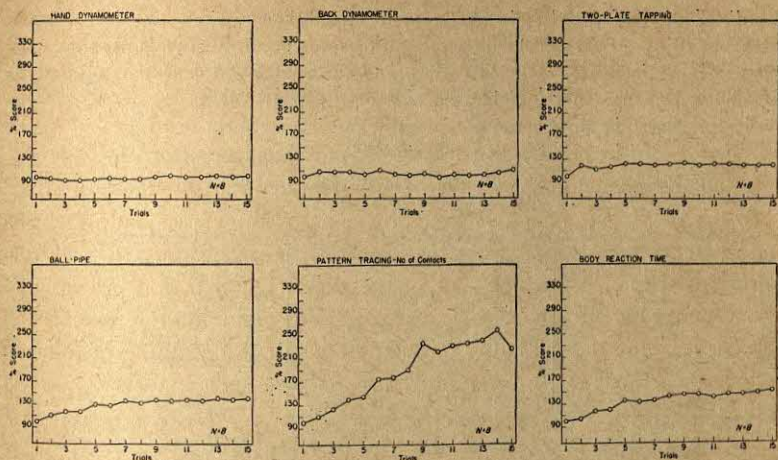


FIG. 1. PRACTICE IMPROVEMENTS IN MOTOR PERFORMANCE
The testing sessions were repeated at intervals of one to two days.

We were interested more in the duration of practice necessary for reaching a plateau rather than in an accurate evaluation of the relative improvement in the different tests; for this purpose modified standard scores would have provided a better unit of measurement.

For several reasons the three series are not directly comparable. There were differences in the scheduling of the practice sessions—but fairly wide differences in

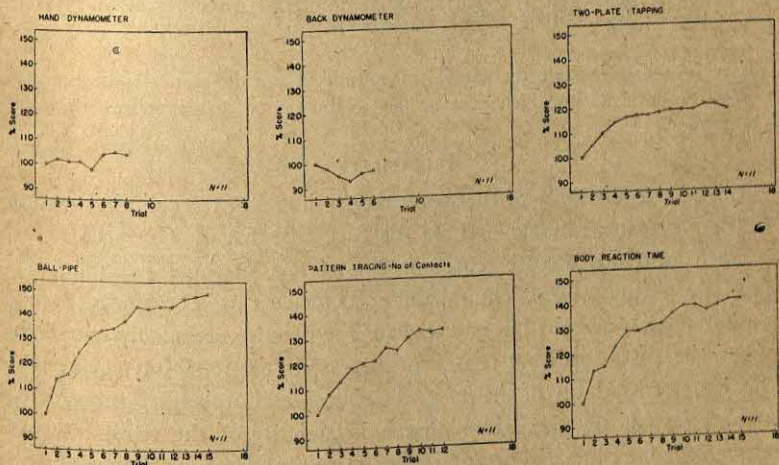


FIG. 2. PRACTICE IMPROVEMENTS IN MOTOR PERFORMANCE
The testing sessions were scheduled at intervals of two to three days.

the distribution of practice do not seem to affect the practice gains,¹⁴ at least as far as coördination (tracing) and reaction-time are concerned. In Series 1 the testing was more intensive, as noted earlier, and there were differences in equipment. 'Good' performance on the eighteenth trial of the Series 3 was heavily stressed in the instructions. The three series are, nevertheless, in agreement that tests of strength are affected least by practice; that tests involving a significant amount of coördination show the longest period of practice; and that the speed of tapping is inter-

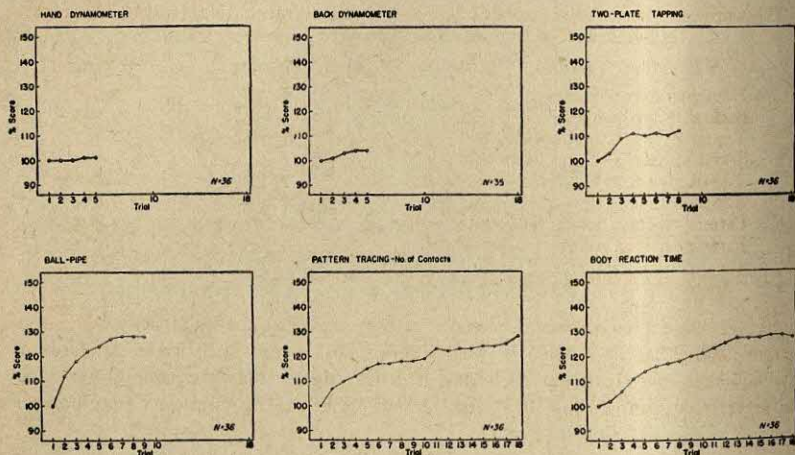


FIG. 3. PRACTICE IMPROVEMENTS IN MOTOR PERFORMANCE

The testing sessions were repeated at different intervals in different Ss (from 3 times a day to 3 times a week, one day apart as well as irregularly).

mediary. The data indicate the general magnitude of practice-effects to be expected in stress experiments in which tests similar to those used in the present study are to be applied.

RELIABILITY

Data on the reliability of the scores on psychomotor tests administered to 36 Ss are presented in Table II. In the back-pull dynamometer the N was 35. Each individual's score represents the average of three determinations made at a testing session. The reliability coefficients were computed for the two initial trials and the last two trials of the practice period. As one might expect, there is a small but consistent increase in the reliability coefficients. This increased stability of the measurements should probably be considered as the effect of practice, although the irregularity of the retest intervals between the initial trials (first and second) spoils a clear-cut argument.

¹⁴ J. C. Franklin and Brožek, The relation between distribution of practice and learning efficiency in psychomotor performance, *J. Exper. Psychol.*, 37, 1947, 16-24.

The reliability of the measurements of 'voluntary' motor functions is fairly satisfactory, both on the basis of absolute criteria and in comparison with measurements of a variety of biochemical and physiological characteristics.¹⁵ The usefulness of the r is, however, limited. In laboratory studies concerned with the impact of varia-

TABLE II

RELIABILITY COEFFICIENTS (TEST-RETEST r_s) FOR PSYCHOMOTOR TESTS AT THE START OF PRACTICE AND AT OR NEAR THE PLATEAU OF PRACTICE

The two initial trials were repeated within a week, at varying intervals for different individuals. The two terminal trials were made a week apart.

Test	At start	At plateau
Handgrip dynamometer	0.795	0.946
Back-pull dynamometer	0.864	0.925
Tapping (paper-and-pencil)	0.873	0.885
Speed of hand and arm movements (ball-pipe test)*	0.758	0.849
Reaction-time	0.864	0.880
Pattern-tracing, duration of contacts	0.871	0.891
Pattern-tracing, number of contacts	0.745	0.834

* Performed in a standing position, not on the treadmill.

tions of external environment, diet or activity on performance-capacity, the number of Ss that can be studied is usually very small, if measured by the standards of psychometric investigations. As a rule it is feasible to study the same group of Ss on several occasions during the control period rather than to measure a large number

TABLE III

COEFFICIENTS OF CONSISTENCY (r_c) FOR MEASUREMENTS REPEATED AT ONE MONTH'S INTERVAL
(12 Ss , 3 trials; SD^2 -values based on $N=36$)

Test	V_R	V_{WD}	r_c^*	SD^2	Corrected r_c
Handgrip dynamometer	4.09	77.15	0.947	68.72	0.940
Back-pull dynamometer	36.67	980.00	0.963	772.84†	0.953
Tapping (on treadmill)	6.50	37.88	0.828		
Tapping (paper-and-pencil)	5.09	19.58	0.741	29.10	0.825
Speed of hand and arm movements (ball-pipe test)	6.55	38.30	0.829	34.32	0.809
Reaction-time	2.36	60.52	0.961	12.72	0.814
Pattern-tracing, duration of contacts	4.31	42.85	0.899	55.94	0.923
Pattern-tracing, number of contacts	5.91	83.55	0.929	177.09	0.967

* $r_c = 1 - (V_R/V_{WD})$, where V_R =random variance, and V_{WD} =variance within days (average interindividual variance).

† $N=32$.

of Ss on only two occasions. Under these conditions the coefficient of product-moment correlation, the classical reliability statistic, is not applicable. It is the relation of random variance (V_R , intra-individual variance corrected for the trend of the mean scores obtained on repeated testing trials) to the inter-individual variance (V_{WD} ,

¹⁵ Brožek, Taylor, and Keys, Variabilité et fidélité des mesures biologiques chez les jeunes gens, *Travail Humain*, 16, 1953, 71-80.

variance within days or trials), in the form of a coefficient of consistency, r_c , which provides a proper statistical tool for characterization of the stability of repeated measurements. It is applicable both to data which exhibit a practice trend and to data obtained after a plateau has been reached.¹⁶

The reliability of repeated measurements is affected by a variety of factors, including the length of the period between the trials, and by the relative homogeneity-heterogeneity of the S s included in the sample. Usually, the two trials (test-retest) are spaced quite closely, with an interval of one or a few days. We had an opportunity to study the consistency of measurements of motor functions over a period of 3 mo., with the measurements made at approximately monthly intervals, after a stable practice plateau had been established. The measurements served as control data for interposed stresses. The reliability data are presented in Table III.

In addition to the coefficients of consistency, r_c , the variances on which the coefficients are based are also given. This provides the opportunity for considering the variability of the small sample ($N = 12$) as compared with a larger sample of individuals drawn from the same population (normal young men, taking part as conscientious objectors in work projects under Civilian Public Service). For most tests the standard deviations obtained for the larger sample are the best available

TABLE IV
INTERCORRELATIONS BETWEEN MOTOR TESTS AT THE START OF PRACTICE
($N=35$; the values of r required for significance at the 5- and 1-% levels are 0.334 and 0.430, respectively.)

Test	(2)	(3)	(4)	(5)	(6)	(7)
(1) Handgrip dynamometer	.708	.345	-.109	.144	.184	.046
(2) Back-pull dynamometer	—	.259	.133	.260	.299	.025
(3) Tapping (paper-and-pencil)		—	.042	.137	.264	.309
(4) Speed of hand and arm movements (ball-pipe test)			—	.041	.340	.481
(5) Reaction-time				—	.184	.157
(6) Pattern-tracing, duration of contacts					—	.513
(7) Pattern-tracing, number of contacts						—

estimates of the scatter of the scores in the population. Interestingly enough, the scatter in the two samples is fairly similar in most of the tests. In the complex reaction-time the scatter is considerably larger in the smaller sample, due to the inclusion of one very slow individual, while the inter-individual variance (SD^2) in the pattern tracing, especially in the number score, is higher in the 36-man sample. The last column of Table III presents coefficients of consistency in which the large-sample variance was substituted for the average inter-individual variance (variance within days, V_{wd}) of the smaller sample. Correction of coefficients of reliability to a standard scatter was recommended by Cattell.¹⁷

The correction does not change the picture dramatically, except for one or two tests. It may be noted that the practice-plateau values of Table II and the coefficient of consistency in Table III are rather similar. In a well-motivated group, practiced to

¹⁶ Brožek and Howard Alexander, A note on the estimation of the components of variation in a two-way table, this JOURNAL, 60, 1947, 629-637.

¹⁷ R. B. Cattell, *A Guide to Mental Testing*, 1948, 5.

a stable plateau and subsequently involved in an intensive testing-program, the mean control measurements made at monthly intervals were surprisingly consistent.

INTERCORRELATIONS

Ideally, one should like to use a battery of tests with low intercorrelations and large 'loadings' on important motor factors. The *rs* have been computed both at the start of practice and at a time when the group reached or closely approached a plateau of performance (Tables IV and V).

For purposes of easier interpretation, the direction of the deviations from the mean score was reversed for three variables (reaction-time and pattern tracing, both scores). In this way a positive deviation had biological meaning, *i.e.* indicated a

TABLE V
INTERCORRELATIONS BETWEEN MOTOR TESTS AT OR NEAR THE PLATEAU OF PRACTICE
(*N*=35)

Test	(2)	(3)	(4)	(5)	(6)	(7)
(1) Handgrip dynamometer	.710	-.063	.040	.045	.126	.123
(2) Back-pull dynamometer	—	.070	.197	.209	.074	.018
(3) Tapping (paper-and-pencil)		—	.136	.197	.393	.467
(4) Speed of hand and arm movements (ball-pipe test)			—	.306	.364	.162
(5) Reaction time				—	.294	.071
(6) Pattern-tracing, duration of contacts					—	.683
(7) Pattern-tracing, number of contacts						—

better performance in all tests. To reduce the effect of the errors of measurement, the score of each *S* was based on two consecutive trials (six determinations).

In general, the motor tests tend to correlate positively with each other but the *rs* are usually low.¹⁸ In the present test-battery the intercorrelations only infrequently reached a level of statistical significance. Only those tests which were functionally very close, such as the two dynamometers or the two scores on the pattern tracing test, had fairly high intercorrelations.

There are no consistent differences in the magnitude of the correlation coefficients obtained at the start and at the end of intensive practice, with mean *r* values of 0.239 and 0.236, respectively. A *z*-transformation of the *rs* was used in arriving at the average figure.¹⁹

SENSITIVITY TO STRESS

For purposes of describing the effects of a particular stress it is desirable to measure as many facets of fitness as are compatible with the experimental schedule and the competence of the investigators. In 'fundamental' studies it makes no difference whether the results of a particular test are positive,

¹⁸ R. H. Seashore, *Experimental and theoretical analysis of fine motor skills*, this JOURNAL, 53, 1940, 86-98.

¹⁹ E. F. Lindquist, *Statistical Analysis in Educational Research*, 1940, 215.

indicating a definite change, or 'negative.' In either case the results contribute to a description of the overall biological impact of a stress. The magnitude of the stresses should preferably be close to the limits of tolerance.

In the context of 'applied' research one is frequently concerned with the effect of different intensities of the stress and, more importantly, with evaluation of the effectiveness of a variety of physiological (diet, acclimatization) and technological (clothing) measures aiming at reducing the stress or increasing stress tolerance. For this purpose, inclusion of measurements which are not sensitive to the particular type of stress represents a sheer waste of time and manpower. Yet the available information is very incomplete.

In evaluating a mean change, we may apply two basically different criteria. Generally only the first criterion is considered, *i.e.* the *statistical significance* of the change, depending on the *consistency* of the changes taking place in the individual Ss, their magnitude, and the size of the sample. For this purpose, the *t*-test for paired variates or the equivalent *F*-test is appropriate.²⁰ The second criterion refers solely to the *magnitude* of the difference between the control and the experimental means, reflecting the *biological importance* of the change. This magnitude is best evaluated by referring the mean change to the scatter of the values in the normal population (SD_P). The displacement ratio is then obtained as $DR = (M_O - M_E)/SD_P$, where M_O is the mean score for the control and M_E for the experimental period. The value is not affected by the size of the sample. The sign attached in the tables to the displacement ratio indicates the biological direction of the change, 'plus' indicating improvement, 'minus' standing for deterioration.

In actual practice the standard deviations of the population (SD_P) are rarely available and the SD_P is replaced by the estimate made on the basis of the data obtained for a given sample in the control period, SD_O , or for some other sample considered more representative of the parent population. In the present study a SD value based on a sample of 34 Ss will be used, when available.

The stresses to be considered here were carried out in two separate groups of Ss and can be classified chronic and acute.

The chronic stress was semi-starvation with 32 Ss completing the experiment. The food intake was decreased from 3500 cal. per day to 1570 cal. per day for a period of 24 weeks. The stress resulted in a decline of body weight from 69.4 kg. to 52.6 kg.²¹ The protein, fat, mineral and vitamin contents of the starvation diet were adequate. In addition to weight loss, the Ss showed the classical signs and symptoms of famine victims; namely, edema, anemia, polyuria, bradycardia, weakness, and depression. During the semi-starvation period, the Ss followed a set schedule of routine laboratory and housekeeping chores, 20 miles of walking each week, and some men took part in special educational activities.

In addition to the semi-starvation, 4 acute stresses were studied on 12 men, none of whom took part in the semi-starvation regimen. The men were resident in the laboratory during the entire period covered by the several stresses. The first stress was preceded by a control period of 2 mo. During this period the men became

²⁰ Brožek and Alexander, The Formula $t^2 = F$, this JOURNAL, 63, 1950, 262-269.

²¹ Keys et al., *op. cit.*, 1950, 703-707.

accustomed to all test-procedures. The stresses were separated by periods of 6 wk. to 2 mo. and were administered in the order in which the conditions are described below:

(1) *Lack of sleep.* The stress consisted of a sleep-deprivation of 62 hr. The physical activity was kept at the accustomed moderate level.

(2) *Heat stress.* Unacclimatized Ss were placed in experimental rooms with temperatures maintained at 115° to 120°F. during the day time and 90°F. at night, with a low humidity (25%). Work consisting of seven 10-min. periods of walking on the motor-driven treadmill at 7.5% grade and 3.5 m.p.h. alternated with 10-min. rest-periods was carried out on each half day of the experiment. Water was allowed *ad libitum*.

(3) *Hard physical work.* The experimental regimen involved an abrupt increase in the amount of aerobic work. This was accomplished by increasing the assigned treadmill-work from 1 hr. a day at 3.5 m.p.h. on a 10% grade to 6 hr. a day for

TABLE VI
HANDGRIP (IN KG.) UNDER VARIOUS CONDITIONS OF STRESS
(SD=8.3, based on N=34)

Stress	N	Mean change	F-test	DR†
Semi-starvation	32	-16.4	214.95*	-1.98
Acute starvation	12	+ 0.5	0.26	+0.06
Hard work	10	- 2.8	36.00*	-0.34
Heat	12	- 2.7	18.53*	-0.33
Lack of sleep	12	+ 0.2	0.05	+0.02

* Significant at 1.0% level.

† Displacement ratio = $(M_C - M_E)/SD$, where M_C =control mean; M_E =experimental mean; SD=estimate of the standard deviation of the population. A plus sign (+)=improvement, a minus sign (-)=deterioration.

2 days and 4 hr. a day for the third day. One half hour of rest was allowed between work-periods. This increased the approximate daily energy expenditure from 3500 to 5800 cal.

(4) *Acute starvation with hard work.* The men ate no food (water *ad libitum*) and walked on the motor-driven treadmill for four hours daily with interposed rest periods. These conditions resulted in a caloric deficit of from 3500 to 4000 cal. a day.²² The men developed dehydration, hypoglycemia, acetonuria and a mild acidosis during the course of the experimental period.²³

The psychomotor tests were given daily in all the short-term stresses. For the present purposes only the maximal changes will be given. These changes represent the terminal period of the stress, the important exception being the heat stress, in which the maximal changes occurred on the evening of the first day. In the semi-starvation experiment the data were obtained in the last 2 wk. of the 24-wk. period of reduced food-intake. The data for individual measurements are given in Tables VI-XII.

In the handgrip the largest change, in terms of either the F-test or the displacement ratio, took place in the semi-starvation experiment. The mean score was dis-

²² Henschel, Taylor, and Keys, Performance capacity in acute starvation with hard work, *J. Appl. Physiol.*, 6, 1954, 624-633.

²³ Taylor, Henschel, Mickelsen, and Keys, Some effects of acute starvation with hard work on body weight, body fluids and metabolism, *J. Appl. Physiol.*, 6, 1954, 613-623.

placed about two standard deviations below the average value obtained during the control period. In the hard work and heat stress the absolute change was small and would not be considered as a practically important impairment of the work capacity. Though small, the changes were fairly consistent from S to S and, consequently, statistically highly significant. This contrast between the two criteria of the 'significance' of mean change is clear enough for textbook purposes. The lack of sleep did not result in a decreased performance on this test. More interesting, however, is the

TABLE VII
BACK-PULL (IN KG.) UNDER VARIOUS CONDITIONS OF STRESS
(SD=27.8, based on N=29)

Stress	N	Mean change	F-test	DR
Semi-starvation	29	-49.7	179.32*	-1.79
Acute starvation	12	-7.2	3.97	-0.26
Hard work	9	-12.2	16.21*	-0.44
Heat	11	-15.3	12.69*	-0.55
Lack of sleep	10	-7.1	7.68†	-0.26

* Significant at 1-% level.

† Significant at 5-% level.

absence of deterioration of strength of grip in acute starvation. This does not appear to be an artifact. When the Ss were exposed to the same stress on another occasion, the mean decrement in the score was 1.9 kg., but the change did not reach a level of statistical significance ($F=1.28$).

The pattern of changes in the back-pull dynamometer, a test disliked by most of the Ss, was fairly similar to strength of grip. The changes were, on the whole, more marked, except in the semi-starvation study. The magnitude of deterioration in the

TABLE VIII
TWO-PLATE TAPPING, NUMBER OF TAPS PER 10 SEC.
(SD=5.4, based on N=34)

Stress	N	Mean change	F-test	DR
Semi-starvation†	32	-3.3	34.10*	-0.61
Acute starvation	12	-10.3	28.04*	-1.91
Hard work	10	-1.7	5.41†	-0.31
Heat	11	-1.7	1.61	-0.31
Lack of sleep	12	-1.0	4.34	-0.19

* Significant at 1-% level.

† Significant at 5-% level.

‡ A paper-and-pencil version of the test used here.

acute starvation and the lack-of-sleep experiments was identical but only the latter change exceeded the 5-% level of statistical significance.

With reference to the tapping test, only the last four stresses are directly comparable as the standard tapping test, performed while walking on the treadmill, was used. This test of speed is highly resistant to deterioration and the changes were very small, except in acute starvation. In semi-starvation, a paper-and-pencil version of the test was used. It is of interest that while the control means and the scatter of the scores on the two versions of the test are comparable, the apparatus-

test is much more sensitive to stress. Thus in another acute starvation experiment, where both versions of the test were used, the group (paper-and-pencil) version showed no deterioration, with mean scores of 72.3 during the control period and 71.8 and 72.3 on the third and fourth day of the experiment. The mean scores on the regular test showed a marked change. The values were 71.0 in the control period and 60.4 at the end of the fourth day of fasting, yielding a decrement of 10.6—a value very close to that given in Table VIII—and an *F*-test of 33.38

TABLE IX

SPEED OF HAND AND ARM MOVEMENTS, NUMBER OF TIMES A BALL IS PASSED
THROUGH A VERTICAL PIPE IN 1 MIN.
(*SD*=5.9, based on *N*=34)

Stress	<i>N</i>	Mean change	<i>F</i> -test	<i>DR</i>
Semi-starvation†	32	- 3.9	31.49*	-0.66
Acute starvation	12	-11.0	16.78*	-1.86
Hard work	10	- 5.2	16.49*	-0.88
Heat	11	- 4.3	4.15	-0.73
Lack of sleep	12	- 3.5	7.19†	-0.59

* Significant at 1% level.

† Significant at 5% level.

‡ Performed standing, not on the treadmill.

(highly significant for 1 and 11 degrees of freedom). In the other three stresses the decrements were either minimal (-0.2 in hard work, -0.5 in heat) or absent (lack of sleep).

The 'ball-pipe' test measuring the speed of coördinated hand and arm movements (Table IX) was also performed, as a rule, on the moving treadmill. The performance on this test was affected more markedly by the biological stresses than the speed of tapping. It was one of the few psychomotor measurements affected by

TABLE X

REACTION-TIME (IN 0.01 SEC.)
(*SD*=3.6, based on *N*=34)

Stress	<i>N</i>	Mean change	<i>F</i> -test	<i>DR</i>
Semi-starvation	32	+3.3	26.59*	-0.92
Acute starvation	12	+9.0	19.80*	-2.50
Hard work	10	+2.1	11.37*	-0.58
Heat	11	+3.1	7.57†	-0.86
Lack of sleep	12	+1.4	5.65†	-0.39

* Significant at 1% level.

† Significant at 5% level.

the lack of sleep. In the semi-starvation experiment, involving a large number of *Ss*, a group-form of the test was used with 6 *Ss* working at a time, standing in front of a bench. This made the test definitely easier—a fact which may explain the relatively small, though consistent, change.

In the whole test-battery, the complex reaction-time was the most sensitive indicator of decreased capacity to work. There was a marked deterioration in acute starvation, with a displacement ratio of 2.50, and the changes were statistically

significant in all five stresses. It will be recalled that the performance involved participation of large muscle groups and some hand-foot coordination. The element of discrimination (three light signals), as such, played probably a minor rôle.

The pattern-tracing was scored in reference to the number and the length of error contacts. The latter score exhibited more consistent changes under stress. It indi-

TABLE XI
PATTERN-TRACING, NUMBER OF CONTACTS
($SD=13.3$, based on $N=34$)

Stress	N	Mean change	F-test	DR
Semi-starvation	32	+20.6	130.42*	-1.55
Acute starvation	12	+9.8	17.12*	-0.73
Hard work	10	+7.6	8.93†	-0.57
Heat	11	+6.4	3.76	-0.48
Lack of sleep	12	+0.1	0.00	0.00

* Significant at 1% level.

† Significant at 5% level.

cated deterioration in the lack-of-sleep stress while the other score remained unchanged; the Ss made the same number of error contacts under control conditions and under stress but it took them a longer time to react to the sound of the buzzer and to get back to the center of the path.

DISCUSSION

Motor performance has been studied in several contexts and for a variety of purposes. Physical education is concerned with such problems as the development of gross motor abilities and the measurement of achievement in athletic activities. Motor fitness represents the capacity to run, jump,

TABLE XII
PATTERN-TRACING, DURATION OF CONTACTS IN $\frac{1}{4}$ SEC.
($SD=7.5$, based on $N=34$)

Stress	N	Mean change	F-test	DR
Semi-starvation	32	+6.1	68.29*	-0.81
Acute starvation	12	+8.5	31.48*	-1.13
Hard work	10	+3.7	4.28	-0.49
Heat	11	+5.6	7.18†	-0.75
Lack of sleep	12	+4.0	9.69*	-0.53

* Significant at 1% level.

† Significant at 5% level.

dodge, fall, climb, swim, lift and carry loads.²⁴ During World War II the Army Air Forces used extensively 'physical fitness ratings' based on a score combining the number of sit-ups, chin-ups, and the duration of a shuttle-run. In South Africa, standards of physical efficiency were based on the time of running 100 and 600 yd., and the putting of a 12-lb shot, con-

²⁴ T. K. Cureton, *Physical Fitness Workbook*, 3rd ed., 1947, 33-42.

sidered as measures of speed, endurance, and strength, respectively.²⁵

The mechanical engineer is concerned with the analysis of cycles of movements into their basic component parts and with effective training methods. Less attention has been paid to the more basic problems of kinesiology.²⁶ On the other hand, some of the recent psychological investigations merge with the mechanical engineer's approach and the method of motion-study.²⁷ In part because of the importance of perceptual elements and the processes of selecting, interpreting, remembering, anticipating, and timing in almost all real life 'motor' performances, the study of skill is a logical meeting ground of the experimental psychologist and the engineer concerned with time-and-motion studies.²⁸

The psychologists have investigated the age-trends in fine motor and mechanical abilities, studied their factorial structure, analyzed the phenomena of motor learning, and used a variety of manipulative tests for the selection and classification of industrial²⁹ and military personnel.³⁰ Tests of motor functions, principally of strength and endurance, were used for an objective evaluation of performance-capacity at different stages of disease and during recovery.³¹ The deterioration of psychomotor performance was investigated under conditions of psychological and physiological stress, such as anoxia, and exposure to heat and cold.³² Nutritional

²⁵ Ernst Jokl, Report of activities of the Medical Research Committee of the National Advisory Council for Physical Education, *Manpower*, 41, 1946 (No. 2), 1-58.

²⁶ C. A. Koepke and L. S. Whitson, Summary of a series of experiments to determine the power and velocity of motions occurring in manual work, *J. Appl. Psychol.*, 25, 1941, 251-264.

²⁷ L. V. Searle and F. V. Taylor, Studies of tracking behavior: Part I. Rate and time characteristics of simple corrective movements, *J. Exper. Psychol.*, 38, 1948, 615-631.

²⁸ R. Conrad, Study of skill by motion and time study and by psychological experiment, *Research*, 4, 1951, 353-358; R. M. Barnes, *Work Methods Training Manual*, 2nd ed., 1947, 75-80; Alphonse Chapanis, W. R. Garner, and C. T. Morgan, *Applied Experimental Psychology: Human Factors in Engineering Design*, 1949, 264-296.

²⁹ W. F. Long and C. H. Lawshe, The effective use of manipulative tests in industry, *Psychol. Bull.*, 44, 1947, 130-148.

³⁰ A. W. Melton, (ed.), *Apparatus Tests*, A.A.F. Aviation Psychol. Progr. Res. Rep. No. 4, 1947.

³¹ Leonard Brahmé, A newly constructed ergodynamometer and its clinical use, *Acta Med Scand.*, 89, 1936, 268-282; R. B. Malmo and H. L. Andrews, A recording device for foot-tapping, with results from polyneuropathic subjects, this JOURNAL, 58, 1945, 247-252; R. S. Schwab, A. L. Watkins, and M. A. B. Brazier, Quantitation of muscular function in cases of poliomyelitis and other motor nerve lesions, *Arch. Neurol. Psychiat.*, 50, 1943, 538-545; H. H. Clarke, Objective strength tests of affected muscle groups involved in orthopedic disabilities, *Res. Quart.*, 19, 1948, 118-147; Improvement of objective strength tests of muscle groups by cable-tension methods, *Res. Quart.*, 21, 1950, 399-419.

³² R. S. Lazarus, James Deese, and S. F. Osler, The effects of psychological stress upon performance, *Psychol. Bull.*, 49, 1952, 295-317; R. B. Malmo and J. L. Finan,

stresses involved such conditions as restricted vitamin intake and partial or total starvation.³³ It is in this framework that the battery of tests described in the present paper was developed and applied.

Descriptions of a large number of psychomotor measurement methods are available in manufacturer's catalogues, in individual papers, and in textbooks.³⁴ A handbook on methods in medical research contains the description of a variety of tests of motor functions especially adapted for laboratory and clinical studies of motor fitness;³⁵ no overall view of the field has been attempted and the individual contributions are largely limited to apparatus and testing procedures. Information on the effects of practice, reliability, inter-relationships under normal conditions and under stress, and sensitivity to stress is inadequate or unavailable.

With reference to laboratory testing of motor fitness, four components or aspects of voluntary motor performance have been distinguished in the past: speed, strength, coördination, and endurance. It is recognized that these are abstract categories of behavior, *i.e.* biologically non-existent divisions of human motor activity.³⁶ Statistically, the unitary character of these categories is weak, as indicated by low co-variation between scores in tests belonging 'logically' to the same category (modality). Thus Seashore, Dudek, and Holtzman noted that intercorrelations among various tests, all of which presumably depend on steadiness and precision of movement, can not be adequately accounted for by postulating a single factor of steadiness.³⁷ Physiologically, any kind of 'psychomotor' response consists of central (nervous excitations) and peripheral processes (conduction of impulses and muscular contraction). The final mechanism controlling muscle action is the same throughout and operates by the variation in the number

A comparative study of eight tests in the decompression chamber, this JOURNAL, 57, 1944, 389-405; Melton, *op. cit.*, 945-980; Henschel Taylor, Brožek, and Keys, Vitamin C and ability to work in hot environments, *Amer. J. Trop. Med.*, 24, 1944, 259-265; N. H. Mackworth, Researches on the measurement of human performance, *Medical Research Council, London*, 1950, Special Rep. Ser. No. 268; *Varii*, The tolerance of man to cold as affected by dietary modifications, *Amer. J. Physiol.*, 146, 1946, 66-83, 84-96, 538-558.

³³ Brožek, Guetzkow, Mickelsen, and Keys, *op. cit.*, 1946, 360; Keys et al. *op. cit.*, 1950, 703-707; Taylor, Brožek, Henschel, Mickelsen, and Keys, The effect of successive fasts on the ability of men to withstand fasting during hard work, *Amer. J. Physiol.*, 143, 1945, 148-155.

³⁴ A. G. Bills, Studying motor functions and efficiency, in T. G. Andrews, ed., *Methods of Psychology*, 1948, 459-497; Henri Piéron, Méthodes expérimentales, in *Méthodologie psychotechnique, Traité de psychologie appliquée*, 1, 1952, 202-210.

³⁵ W. R. Miles, ed., Selected psychomotor measurement methods, *Methods in Medical Research*, 3, 1950, 142-218.

³⁶ R. B. Cattell, *Description and Measurement of Personality*, 1946, 72.

³⁷ R. H. Seashore, F. J. Dudek, and Wayne Holtzman, A factorial analysis of arm-hand precision tests, *J. Appl. Psychol.*, 33, 1949, 579-584.

of active motor units in a particular muscle and the frequency of nerve impulses activating these units (the frequency of discharge).³⁸ The physiological differentiation of motor functions depends on the relative importance and function of the peripheral (e.g. muscle size) and central components, the participation of different senses and different segments of the central nervous system, and the degree of involvement of other physiological systems (circulation, respiration). It becomes readily apparent that the four categories of motor functions do not represent, physiologically, homogeneous entities.³⁹ Thus in the 'speed' of small repetitive movements (e.g. tapping) the limiting factor appears to be the refractory period of the motor centers while the speed of single ballistic movements reflects the force developed during isotonic contraction and the speed of simple reaction (reaction-time) is principally a measure of the delay at the central sensori-motor synapses. 'Endurance' in *static* work, such as weight holding, involves tolerance of discomfort resulting from peripheral ischemia while in severe *dynamic work*, such as running on a motor-driven treadmill, the cardio-respiratory system (heart output, pulmonary capacity) and the ability to accumulate oxygen debt are the crucial factors. In two-plate tapping, the rate of tapping falls off as a result of disturbed function of the motor centers (including 'blocking') rather than the inability to contract the muscles involved in this simple task. 'Coördination' is also not a unitary category involving, as it does, simultaneous contraction (and relaxation) of several muscle groups, coördination of actions in temporal sequence (timing) as well as coördination of sensory perception and neuromuscular function. It has been noted that the degree of recovery after experimental starvation has grouped a test of the 'speed' of ballistic leg movements together with the measures of strength rather than with the tests of speed as measured by tapping or the time of a complex reaction.⁴⁰

Statistically, in terms of factorial analysis, there is weak or no evidence for a general motor factor.⁴¹ The group factors are narrow. In the category of 'speed,' speed of single reactions does not correlate highly with the speed of tapping or with serial discriminative reactions.⁴² This might be expected from the physiological analysis presented above. In the present

³⁸ Ernst Gellhorn, *Physiological Foundations of Neurology and Psychiatry*, 1953, 15-18.

³⁹ Simonson, Endocrinologic aspects of aging and the muscular and nervous system in man, *J. Gerontol.*, 7, 1952, 303-305.

⁴⁰ Brožek, Simonson, and Keys, *op. cit.*, 1952, 757.

⁴¹ H. G. Seashore, Some relationships of fine and gross motor abilities, *Res. Quart.*, 13, 1942, 259-274.

⁴² Seashore, C. E. Buxton, and I. N. McCollom, Multiple factorial analysis of fine motor skills, this JOURNAL, 53, 1940, 251-259.

study the three tests of 'speed' had very low intercorrelations while the speed of hand and arm movements correlated fairly highly with the test of eye-hand (foot) coördination.

The narrowness of group factors makes difficult a comprehensive evaluation of the motor aspects of the overall fitness as the number of tests which can be incorporated into an experimental program of interdisciplinary research is limited. Perhaps a heroic attack on the factorial structure of motor functions, similar to Thurstone's studies of intelligence,⁴³ would provide the applied psychologist and physiologist with a better guidance as to the selection of an economic battery of motor tests than has been available in the past. Both a comprehensive coverage and the use of well-constructed tests is a prerequisite for a fairly definitive solution of this problem.

For the use of motor tests in longitudinal experimental investigations it is fortunate that practice on tests of motor skills does not seem to raise significantly the intercorrelations between scores on tests of fine motor skills.⁴⁴ In the present study, the mean r values at the start and end of intensive practice were practically identical. Were this not the case the factorial analysis would have to be done on scores obtained at the practice plateau—an almost hopeless prospect.

Factorial studies made on scores obtained under 'normal' (control) conditions cannot possibly, however, answer the crucial question of the relative sensitivity of different functions to the impact of biological stresses. In a very real way, the study of the pattern of changes in motor functions under stress represents a physiological factor analysis of motor fitness. The available data indicate, again, a good deal of independence, especially as far as the relative *magnitude* of deterioration is concerned. Thus in acute starvation strength was not significantly affected while all three speed-tests showed a very marked deterioration; coördination showed a definite but smaller change, using the displacement ratio as criterion of the relative magnitude of the change. In acute thiamine deficiency all motor functions exhibited a tendency toward deterioration.⁴⁵ While the direction of the change was shared by all tests, strength of grip and speed of tapping changed but little, the speed of hand and arm movements and the reaction-time were definitely affected, and performance on the pattern-tracing test of coördination deteriorated dramatically.

⁴³ L. L. Thurstone and T. G. Thurstone, *Factorial Studies of Intelligence*, 1941, 1-94.

⁴⁴ Buxton and L. G. Humphreys, The effect of practice upon intercorrelations in motor skills, *Science*, 81, 1935, 441-442.

⁴⁵ Brožek, Guetzkow, Mickelsen, and Keys, *op. cit.*, 1946, 376.

This battery of motor tests has been used as part of a larger battery of tests designed to cover several aspects of fitness.⁴⁶ It has been pointed out that motor functions are in general more sensitive to stress than indices of intellectual capacity and personality patterns examined by projective tests.⁴⁷ Special complaint inventories and some personality questionnaires proved more useful.⁴⁸ The relationship between the physiological indices of fitness and the motor-tests described here are of some interest. In studying the response to stress, the physiologist has placed emphasis (1) on indices of cardiovascular function—the rate of the pulse at rest and at work, the response of pulse and blood pressure to posture, the maximal oxygen intake; (2) on indices of the capacity to do anaerobic work—the concentration of lactate or pyruvate in the blood and the oxygen debt after exhausting exercise; and finally (3) on the ability to perform exhausting exercise (high energy expenditure), usually measured by time required to run to exhaustion. In more elaborate studies, the physiologist has measured skill by determining the mechanical efficiency of work and has used the ability to walk on the treadmill (at 3.5 m.p.h. on a 10% grade) for long periods as an index of 'endurance' in carrying on activity which raises only moderately (6-7 times) the resting energy expenditure. The concentration of physiological measurements on the cardiovascular system and on performance in exhausting exercise is due to the recognition of the fact that cardiovascular reflexes are markedly influenced by both physical conditioning and deconditioning processes and that exhausting exercise is the result of such a large number of integrated physiological and psychological processes that this test-situation is likely to be very sensitive to changes in the status of the individual.

On a *a priori* grounds one cannot be certain whether a battery of motor-tests such as described here would be more sensitive to certain stress-situations than the usual measurements employed by the physiologist. Actual experience demonstrates that tests of motor performance are more sensitive

⁴⁶ Keys, Henschel, Taylor, Mickelsen, and Brožek, Experimental studies on man with a restricted intake of the B vitamins, *Amer. J. Physiol.*, 144, 1945, 5-42.

⁴⁷ Franklin and Brožek, The Rosenzweig P-F test as a measure of frustration response in semistarvation, *J. Consult. Psychol.*, 13, 1949, 293-301; Brožek, Guetzkow, M. V. Baldwin, and Robert Cranston, A quantitative study of perception and association in experimental semistarvation, *J. Personality*, 19, 1951, 245-264; N. K. Kjenaas and Brožek, Personality in experimental starvation, *Psychosom. Med.*, 14, 1952, 115-128.

⁴⁸ Brožek and B. C. Schiele, Clinical significance of the Minnesota Multiphasic F Scale evaluated in experimental neurosis, *Amer. J. Psychiat.*, 105, 1948, 259-266; Brožek and N. K. Erickson, Item analysis of the psychoneurotic scales of the Minnesota Multiphasic Personality Inventory in experimental semistarvation, *J. Consult. Psychol.*, 12, 1948, 403-411.

to certain biological stresses than the physiological tests and in others the reverse holds true. Acute thiamine deficiency is an example of the superiority of the motor battery over the physiological measurements,⁴⁹ while bed rest is an example of a stress in which cardiovascular indices showed marked deterioration in the presence of minimal changes in the motor functions.⁵⁰ Endurance in moderately hard physical work, as measured by the ability to complete prescribed periods of walking on the treadmill, was a sensitive indicator of deterioration in the acute thiamine stress.

The psychomotor battery described in this paper does not contain a test of endurance. While the difficulties in constructing adequate tests of endurance are considerable, such tests are of interest because they may serve as more sensitive indicators of experimentally induced deterioration and of clinical impairment than the best performance-tests. Some clinical evidence on this point has been provided.⁵¹ In a patient with mild poliomyelitis, the hip adductors of the right side appeared clinically to be more affected than the corresponding muscles of the left side. When strength of these muscles was tested in single determinations, identical scores were obtained for the two sides. When the measurements were repeated (eight times at intervals of 20 sec.), the performance of the right (weak) side declined rapidly and markedly (from 17 to 10 lb.) while the performance of the left (strong) hip adductors remained essentially constant.

SUMMARY AND CONCLUSIONS

A battery of tests of motor functions, applied in the context of comprehensive studies of 'fitness' and of changes in fitness under a variety of biological stresses, was described. It contained two tests of strength, three of speed, and one of coördination, yielding two scores.

The tests of strength are least affected by practice consisting in repeating the measurements, while tests with a large component of coördination show prolonged practice-trends. In general, 3-6 wk. should be reserved for pre-experimental testing to avoid the complication of the deterioration during stress by continued improvement due to practice.

The consistency of the scores is satisfactory. With practice there is a small but definite increase in the stability of individual repeated measurements.

⁴⁹ Keys et al., *op. cit.*, 1945, 5-42.

⁵⁰ Taylor, Henschel, Brožek and Keys, Effects of bed rest on cardiovascular function and work performance, *J. Appl. Physiol.*, 2, 1949, 223-239.

⁵¹ C. O. Molander and Bertha Weinman, A repetitive resistance test for muscular fatigue, *Arch. Phys. Ther.*, 23, 1942, 276-280.

Practice does not consistently alter the amount of correlation between motor-tests. The coefficients of correlation are overwhelmingly positive but they are low except for tests which measure very similar aspects of performance capacity, such as handgrip and back-pull tests of strength.

It has been our general experience that motor performance is more susceptible to deterioration under stress than the sensory and intellective functions. Even in the motor area the changes have frequently been small under conditions of severe stress. The complex reaction-time was the only variable which exhibited statistically significant changes in all five stresses (semi-starvation, acute starvation, hard work, heat, and lack of sleep). Strength showed greatest decrement in semi-starvation while speed of tapping, speed of hand and arm movements, and complex reaction-time exhibited largest deterioration in acute starvation combined with hard physical work.

Examples are quoted to show that the combination of a battery of tests of motor function plus physiological indices of the ability to do hard physical work measures deterioration in the presence of stress better than either approach used alone and provides a more meaningful basis for analysis of fitness deterioration.

In addition to the criterion of statistical significance, a 'displacement ratio' (mean change under stress to standard deviation of the scores in the control period) was used to evaluate the biological importance of the decrements in individual motor functions.

PERFORMANCE VARIABLES IN THE EXPERIMENTAL ANALYSIS OF THE LAW OF EFFECT

By LEO POSTMAN and PAULINE AUSTIN ADAMS, University of California

Thorndike's law of effect is both a law of learning and a law of performance. If a stimulus-response sequence is followed by a satisfier, the connection between stimulus and response is strengthened and the probability that the response will be given to the stimulus is increased.¹ Neither the formulation of the law nor the experimental procedures used to test the law allow of a separation between the conditions of acquisition and the conditions of performance of a habit. Such a distinction is, of course, central in the learning theories of both Hull² and Tolman.³ Both theories recognize that the laws governing the growth of habit strength, or the formation of expectancies, are not sufficient for the prediction of concrete performances in specific test-situations. The activation of habits or the utilization of bits of knowledge acquired during training depends on special principles of performance. This conceptual distinction between learning and performance has as yet received little application to the analysis of the results obtained in typical Thorndikian situations, such as those using the selective reinforcement of word-number connections. There has, to be sure, been considerable discussion concerning the proper interpretation of the fact that announcement of 'Right' following a word-number sequence increases the probability of that sequence on later trials.⁴ Does the announcement of 'Right' stamp in a connection between word and number or does it provide the S with information which he uses on later trials if and when he is motivated to repeat the reinforced response?

An important attempt to distinguish between these two interpretations of effect

* Accepted for publication October 19, 1953. This research was facilitated by a grant from the Behavioral Sciences Division of the Ford Foundation.

¹ The most comprehensive survey of experimental evidence supporting the law of effect may be found in E. L. Thorndike, *The Fundamentals of Learning*, 1932, 170-327.

² C. L. Hull, *Principles of Behavior*, 1943, especially ch. 14.

³ E. C. Tolman, *Purposive Behavior in Animals and Men*, 1932, especially ch. 22. See also Tolman, The law of effect: a round table discussion: II, *Psychol. Rev.*, 45, 1938, 200-203.

⁴ For a review of the different interpretations of the law of effect, see Leo Postman, The history and present status of the law of effects, *Psychol. Bull.*, 44, 1947, 489-563.

was made in a series of experiments by Wallach and Henle.⁵ These investigators informed their Ss that their experiment tested success in extra-sensory perception, hence numbers called correct on one trial might or might not be correct on the next trial. These instructions, according to Wallach and Henle, served to eliminate the 'learning motive' and rendered the rewards irrelevant to the Ss' task. Under these conditions, responses called 'Right' were not repeated more often than responses called 'Wrong,' nor was the level of repetition of wrong responses as high as that typically obtained in Thorndike's experiments. With elimination of the learning motive, rewards no longer had the automatic effect claimed for them by Thorndike. This analysis fails to distinguish clearly between the conditions governing learning and the conditions governing performance. The instructions given at the beginning of an experiment may influence both learning and performance, or performance alone. The announcements of 'Right' may produce learning which does not manifest itself in performance when Ss are not motivated to use the information they have acquired.

To separate the effects of motivation on learning and on performance, we must sharply distinguish between two questions: (1) Does S learn though he is not motivated to do so, *i.e.* does he acquire information about the correct word-number connections even though that information is irrelevant to his task? (2) If learning does occur, do different instructions regarding repetition lead to significant differences in performance? The experiment reported in this paper seeks to answer these questions experimentally.

METHOD

General procedure. A procedure typical of Thorndike's experiments was used. S was presented with a series of 25 three-letter words and instructed to respond to each word with a number from 1 to 10.⁶ All the stimulus-words were common three-letter nouns with a high frequency of usage according to the Thorndike-Lorge word-count.⁷ S's responses to the words in positions 6, 13, and 20 were arbitrarily called 'Right,' all other responses were called 'Wrong.' The series of words was then repeated, but in a different order. The first presentation of the words constitutes the training series, the second presentation the test-series. No rewards or punishments were given for responses in the test-series.

In changing the order of stimulus-words from training series to test-series, the following restrictions, suggested by Jenkins and Cunningham,⁸ were imposed: (a)

⁵ Hans Wallach and Mary Henle, An experimental analysis of the law of effect, *J. Exper. Psychol.*, 28, 1941, 340-349; A further study of the function of reward, *ibid.*, 30, 1942, 147-160.

⁶ Only the responses to 21 of these words were used in the analysis of the results. The first two and last two items were discarded in order to reduce the effects of primacy and recency.

⁷ Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944.

⁸ W. O. Jenkins and Leta Cunningham, The guessing-sequence hypothesis, the 'spread of effect,' and number-guessing habits, *J. Exper. Psychol.*, 39, 1949, 158-168.

No two words followed each other on successive trials. (b) No punished response occurred in the same third of the list on successive trials. (c) Punished responses which preceded a reward on one trial followed a reward on the next trial and vice versa. (d) The positions of the rewarded responses were held constant. To reduce the possible effects of specific stimulus-words on repetition, two different orders of the list were used in the training series with different words in the rewarded positions. The two different lists were used equally often under all experimental conditions.

Following the test-series, a recall-test was given in which *S* was instructed to recall the numbers with which he had responded during the training series and the reinforcements he had received. For the recall-test, the words were given in a new order, obeying the same restrictions as before.

Since the instructions given before the training series and the test-series were a crucial variable in this experiment, they will now be reported in detail.

Training instructions. The essential features of the experimental design were (1) to provide *S* with an opportunity to acquire information about right and wrong word-number connections in the absence of motivation to learn, and (2) to vary systematically *S*'s motivation to use this information on later trials.

To expose *S* to right and wrong word-number connections, without motivating him to learn these associations, we borrowed the procedure of Wallach and Henle and introduced the experiment as a test of extra-sensory perception. The following instructions were given *S* prior to the training series.⁹

You have perhaps heard of extra-sensory perception. Subjects are asked to name items which they cannot possibly perceive in any ordinary sense of the word. If they are able to make hits more frequently than they would by chance alone, it must be by extra-sensory perception.

We are dealing with the same problem by a method somewhat different from the ones used before. We have prepared a list of pairs of words and numbers. I shall show you the words, and you guess the numbers that go with them. If extra-sensory perception exists, you should be able to do this better than chance. If it does not exist, you will merely be guessing the number that goes with each word, and you should not make a score better than chance.

We have used numbers from 1 through 10. So you can choose from 10 numbers in each case, and with this large selection you will, of course, not make many hits. After each judgment I will tell you if you were right or wrong.

Thinking about your performance during the experiment definitely will not help you, so just try to give your responses as quickly as you can. Don't mind if I hurry you through the experiment, it will only help you to make a better score.

We have a list of 25 words, and this list will be presented a number of times, but the numbers assigned to the words will change in a random fashion from presentation to presentation. Thus a number which is correct on the first presentation may or may not be correct on the next.

Are there any questions?

The stimulus-words, typed in capital letters on index cards, were presented at the rate of approximately one every 5 sec. *E* recorded *S*'s response on a record sheet which was not visible to *S*.

Test-instructions. To gauge the effect of differences in motivation on *S*'s performance, with learning held constant, the *Ss* were subdivided into nine experimental

⁹ The instructions used are identical with those of Wallach and Henle, *op. cit.*, 1941, 342, except for a few minor changes in wording.

groups. Prior to the test-series, each of these groups received different instructions which were designed to manipulate *S*'s motivation to repeat or not to repeat responses rewarded or punished during the training series. Different groups were informed that, for the test-series, responses rewarded during training (a) would be correct, (b) might or might not be correct, (c) would be incorrect. Corresponding instructions were given concerning the correctness of the responses punished during the training series.

There are nine possible combinations of such instructions concerning the correctness for the test-series of the responses rewarded and punished during the training series. These combinations of instructions may be seen in Table I. To facilitate reference to the different conditions, the instructions concerning the correctness of responses rewarded or punished during training will be denoted by subscripts: a plus sign indicates that the particular class of responses will be correct for the test-series. Thus R_+ means that previously rewarded responses will be correct, W_+ means that previously punished responses will be correct. Similarly, a minus sign indicates that the class of responses will be incorrect, *i.e.* R_- or W_- . A subscript of zero denotes that the class of responses may or may not be correct, *i.e.* R_0 or W_0 .

Each of the nine experimental groups received one of these combinations of instructions. There were 20 *Ss* in each experimental group.

The instructions for the test-series may be exemplified by those given to members of Group II, which were as follows:

Now we will go through the list again and I will give you the same words as I gave you last time. This time, however, I have chosen the numbers in such a way that the numbers I called 'Right' the first time will be right again this time. The numbers I called 'Wrong' the first time may or may not be wrong this time. This time I will not tell you if your guess is right or wrong.

Comparable instructions were given to the other groups, except that the statements concerning the correctness of the previously rewarded and punished responses varied from group to group.

Recall-test. Immediately after the test-series, a recall-test was given. The purpose of this test was to determine the extent to which repetition of responses during the test-series was associated with active recall of these responses. *S* was also required to state whether or not his responses had been called 'Right' or 'Wrong' during the training series. The recall-instructions read as follows:

Now I want you do one more thing. I will again show you the words. This time I want you to tell me what number you remember saying to each of the words the first time we went through the cards. Please tell me also whether I called the number 'Right' or 'Wrong.' Try to answer as quickly as possible; if you are not sure you can remember, just make a guess.

That is, I want you to give me the same number you gave to each of the cards the first time we went through the list, and tell me whether I called the number 'right' or 'wrong.'

At the end of the experimental session, *E* questioned the *Ss* concerning their interpretation of the instructions. The few *Ss* who gave clear evidence of having misunderstood the instructions were not included in the experiment and were replaced by other *Ss*.¹⁰

¹⁰ Only 3 *Ss* had to be discarded for this reason. *Ss* were also discarded if they

Recall-control group. S's recall of his responses may have been influenced by his performance during the test-series. To obtain a measure of recall independent of the practice provided by the test-series, a tenth group of 40 Ss, to which we will refer as the Recall-Control Group, was added. This group was given a recall-test immediately after the training series. The words were presented in the same order as on the recall-test of the other experimental groups.

Subjects. The Ss were 220 students at the University of California, Berkeley. None knew the purpose of the experiment.

RESULTS

Repetition of rewarded and punished responses. Table I shows for the various conditions the average number of rewarded responses repeated during the test-series. Table II shows the average frequencies of repetition of punished responses. To facilitate comparisons among conditions, the data are presented in 3×3 tables, with the instructions for rewarded responses arranged along the top, and instructions for punished responses

TABLE I
AVERAGE NUMBER OF REWARDED RESPONSES REPEATED BY
THE VARIOUS EXPERIMENTAL GROUPS

The Roman numerals refer to the Experimental Group of Ss which received that particular combination of instructions.

Instructions for

Punished responses	Rewarded responses					
	R_+		R_0		R_-	Totals
W_+	I 1.25	IV 1.20	VII .95			1.13
W_0	II 1.15	V .80	VIII .65			.87
W_-	III .50	VI .55	IX .40			.48
Totals	.97	.85	.67			.83

along the side. Each of the nine cells thus represents one combination of instructions given to one of the experimental groups. The averages for columns make it possible to evaluate the effects of performance-instructions for rewarded responses, summing over instructions for punished responses. Similarly, row averages allow an evaluation of the performance instructions for punished responses, summing over instructions for rewarded responses.

Consider the repetition of rewarded responses (Table I). We may inquire first for which of the conditions the frequency of repetition exceeds the chance level. In answering this question, Thorndike assumed that with

persisted in giving the same number or a very small range of numbers to the stimulus-words, or if they deliberately attempted to make their responses in terms of meaningful associations between words and numbers. Altogether, less than 5% of the Ss had to be replaced for any of these reasons.

ten numbers to choose from the chance probability of a repetition is 0.10 and determined whether the obtained frequency of repetition significantly deviated from that proportion. Although this procedure is open to criticism,¹¹ we shall apply it here to facilitate comparison of our results with those of earlier studies. By chance alone, 6 of the 60 rewarded responses (3 each for 20 Ss) should be repeated. To exceed this chance probability at the 5-% level, 9.82 or an average of 0.49 repetitions per S, are required. Inspection of Table I shows that all but one of the conditions yield repetitions significantly above chance.¹² The sole exception is Condition IX for which the performance instructions are entirely negative. We note that significant levels of repetition are obtained without instruction to learn and under performance instructions which explicitly militate against such repetitions. For the conditions showing the higher numbers of repetitions, the levels compare favorably with those obtained by Thorndike.¹³ Our results do not conform to the conclusion of Wallach and Henle that rewards are effective only when there is a learning motive which makes repetition 'sensible.'

Even though rewards are in general effective, there are important differences among conditions in the frequency of repetition of rewarded responses. Since the conditions of learning were exactly identical for all the experimental groups, the differences in frequency of repetition must be ascribed to the performance instructions.

Comparison of the column averages in Table I shows that positive instructions concerning the correctness of rewarded responses (R_+) lead to a higher level of repetition than do indeterminate instructions (R_0) or negative instructions (R_-). In general the more positive the instructions, the higher is the level of repetition. We note, however, that the differences between column averages are not large, and analysis of variance shows them not to be statistically significant ($F = 2.28$, $df = 2$ and 171 , $P > .05$). One

¹¹ The assumption that the chance level of repetition is 0.10 has been questioned on the basis of the existence of number-guessing habits (cf. Jenkins and Cunningham, *op. cit.*, 164-168). The chance level of repetition is difficult to establish empirically since the absence of either negative or positive reinforcement may be interpreted in various unknown ways by the S. In any event, our main interest is in comparing levels of repetition with instructions to learn (Thorndike, *An Experimental Study of Rewards*, 1933, 16-25), and without instructions to learn (this experiment). Whatever distortion is introduced by assumptions about the chance level hold equally true for both conditions.

¹² In setting this level, the one-tail test of significance was used since the direction of the deviation from chance was predicted. If the two-tail test were used, an average of 0.53 repetitions would be required for the 5-% level. Application of this latter test would leave our conclusions virtually unchanged.

¹³ These levels of repetitions vary around 50% (see Thorndike, *op. cit.*, 4).

might be tempted to conclude that the repetition of rewarded responses is not affected by performance-instructions. Further analysis shows, however, that such is not the case.

Examination of the row-averages of Table I reveals that the more positive the instructions for repetition of *punished* responses, the higher the frequency of repetition of rewarded responses. This relationship is highly significant ($F = 10.67$, $df = 2$ and 171 , $P < .001$). We ascribe this finding to the operation of a general *repetitional set*. The degree of this general repetition set is determined by the nature of the performance-instructions for the punished responses. Punished responses comprise, of course, the great majority of the responses given during the training series. It is plausible, therefore, that *S*'s general readiness to repeat responses given during the training series is influenced by the instructions he is given for the punished responses. To the extent that *S* is set to repeat punished responses, he is ready to repeat most of his responses, including willy-nilly rewarded responses.

We now turn to a consideration of punished responses (Table II). Testing for significant deviation from a chance level of repetition, we

TABLE II
AVERAGE NUMBER OF PUNISHED RESPONSES REPEATED BY THE VARIOUS
EXPERIMENTAL GROUPS
Instructions for

Punished responses	Rewarded responses					
	R_+		R_0		R_-	Totals
W_+	I 5.75	IV	5.25	VII	5.30	5.43
W_0	II 4.05	V	2.40	VIII	3.45	3.30
W_-	III 2.25	VI	3.30	IX	2.65	2.73
Totals	4.02		3.65		3.80	3.82

find that for all the conditions punished responses are repeated above chance at the 5-% level or better.¹⁴ This finding agrees with the results obtained by Thorndike but not with those of Wallach and Henle.

Table II gives additional strong evidence for the importance of the general repetitional set. The more positive the instructions for repetitions of

¹⁴ In determining the number of repetitions significantly different from chance for punished responses, the same procedure was used as in the case of rewarded responses. By the one-tail test, 2.27 repetitions are required. Condition III barely satisfies this criterion with 2.25 responses. If the two-tail test is used, an average of 2.36 repetitions is required. The one-tail test is appropriate if significant repetition of punished responses is predicted on the basis of the law of exercise.

punished responses, the higher is the level of repetition of such responses. It is interesting to note that indeterminate instructions (W_0) act much more like negative than positive instructions. The variance due to instructions for punished responses is highly significant ($F = 13.60$, $df = 2$ and 171 , $P < .001$). Examination of the column averages of Table II shows that the frequency of repetitions of punished responses is independent of the performance-instructions for rewarded responses.¹⁵

In general, then, rewarded responses are repeated at a higher than chance level. The level of repetition is, however, importantly influenced by performance-instructions. The critical instructions are those for punished responses, *i.e.*, the instructions which refer to the majority of the responses in the list. Punished responses too are repeated at a higher than chance level. Again the level of repetition varies significantly with the instructions. Since instructions for punished responses influence the repetition of all responses in the same way, we conclude that the action of rewards and punishments is modified by the operation of a general repetition set.

Differential effectiveness of rewards and punishments. Regardless of the overall level of repetition of responses within a given condition, is the proportion of repetitions higher for rewarded than for punished responses? (It will be remembered that both types of responses tended to be repeated at a better than chance level.) That is, do our data present evidence for the differential effectiveness of rewards and punishments? To answer these questions, the weighted mean difference between the repetitions of the two classes of responses was determined for each of the conditions. The weighting consisted of multiplying the number of rewarded responses repeated by six since there were one-sixth as many opportunities to repeat rewarded responses as there were to repeat punished ones. Table III lists the weighted mean differences for the various conditions and indicates their significance. For all conditions combined, rewarded responses are repeated more frequently than punished ones at a high level of significance ($P < .001$). In the absence of instructions to learn during the original training, rewards were significantly more effective than punishments, just as they are under conditions of intentional learning.¹⁶

The differential effectiveness of reward varies, however, from condition to condition. Rewarded responses are repeated significantly more often than punished ones if, and only if, neither of the performance-instructions is

¹⁵ There was no significant interaction between the instructions for punished responses and the instructions for rewarded responses. The same is true for the results presented in Table I.

¹⁶ Thorndike, *op. cit.*, 1933, 200.

negative (Conditions I, II, IV, and V). If either of the performance-instructions is negative (*i.e.* for all the remaining conditions), rewards and punishments have no differential effects. The difference between these two groups of conditions is highly significant ($F = 9.30$, $df = 1$ and 178 , $P < .001$).

A literal application of Thorndike's hypothesis concerning the automatic effects of satisfiers would demand that rewards be equally and significantly more effective than punishments for all conditions. We say equally effective because the conditions of reinforcement during training were identical for all conditions. Thorndike had, however, recognized that the automatic action of rewards may be modified by *S*'s attitudes and sets.¹⁷

Our finding is that rewards are differentially effective if (a) there is no explicit instruction against repetition of rewarded responses and (b) if the instructions for punished responses do not produce a set against repetition.

TABLE III
WEIGHTED MEAN DIFFERENCES BETWEEN THE AVERAGE NUMBER OF REWARDED
AND PUNISHED RESPONSES REPEATED BY THE VARIOUS EXPERIMENTAL GROUPS
Instructions for

Punished responses	Rewarded responses					
		R_+		R_0		R_-
W_+	I	1.75*	IV	1.95*	VII	.40
W_0	II	2.85†	V	2.40†	VIII	.45
W_-	III	.75	VI	0	IX	-.25

* Significant at the 5% level.

† Significant at the 1% level.

We will now attempt a tentative interpretation of this finding. Let us assume that rewarded associations are, on the average, stronger than punished ones but that the distributions of associative strengths overlap. To demonstrate this difference in associative strength a certain minimal number of the two types of responses must be sampled. It is only when neither performance-instruction is negative that a sufficient number of responses is sampled to bring out the average difference in the associative strength of rewarded and punished responses.

Recall of rewarded and punished responses. The experimental groups were tested for recall of their original responses immediately following the test-series and also attempted to identify the rewards and punishments. The Recall-Control Group was given the recall-test immediately after the training series. The average numbers of rewarded and punished responses

¹⁷ Thorndike, *op. cit.*, 1932, 114 f.

recalled by the various groups are presented in Table IV. There are no significant differences among individual conditions or conditions grouped according to performance-instructions in the number of responses recalled. The experimental groups do not differ significantly from the Recall-Control Group.¹⁸ This lack of significant differences is found for rewarded responses, punished responses and total recalls. *Whereas the frequency of repetitions varies significantly with performance-instructions, recall does not.* Clearly, frequency of repetition and recall are not equivalent tests of the effects of rewards and punishments. There are important differences between the test-series and the recall-situation which may account for this lack of equivalence. As far as the S's task is concerned, the test-series is in every respect identical with the training series; he responds directly to

TABLE IV
AVERAGE NUMBER OF RESPONSES RECALLED BY THE DIFFERENT EXPERIMENTAL GROUPS AND THE RECALL-CONTROL GROUP

Experimental Groups	Instruction	Average number recalled		
		rewarded	punished	total
I	R ₊ W ₊	1.05	5.40	6.45
II	R ₊ W ₀	1.20	6.00	7.20
III	R ₊ W ₋	.85	4.00	4.85
IV	R ₀ W ₊	1.15	5.40	6.55
V	R ₀ W ₀	1.30	4.15	5.45
VI	R ₀ W ₋	1.00	4.60	5.60
VII	R ₋ W ₊	1.30	6.30	7.60
VIII	R ₋ W ₀	1.30	6.60	7.90
IX	R ₋ W ₋	1.00	5.00	5.80
All		1.13	5.27	6.40
Recall Control		1.10	5.35	6.45

the stimulus-words. In the recall-test, S's task is altogether different: he does not respond to the stimulus-word directly but rather attempts to reconstruct his own past responses as well as the reinforcements given by E. The performance-instructions are relevant to what S must do during the test-series but are not relevant to the recall-task. It is not surprising, then, that performance-instructions influence repetitions but no recall.¹⁹ We are not asserting, of course, that repetitions and recalls are altogether independent of each other. The essential point here is that only repetitions are significantly influenced by performance-instructions.

¹⁸ The interpolation of the test-series between training series and recall-test may have provided opportunities both for rehearsal and for retroactive interference. The net effect appears to be negligible.

¹⁹ There were no differences among conditions in the number of intrusions from test-series to recall-test even though there were considerable differences in the number of repetitions in the test-series. This finding lends indirect support to our inference that Ss reacted differentially to the test-series and the recall-test.

Differential recall for rewarded and punished responses. Are rewarded responses recalled relatively more frequently than punished ones? Summing all conditions, we find that an average of 1.13 or 37.7% of rewarded responses are recalled, and 5.27 or 29.3% of the punished responses. Using the same weighting procedure as before to equate the two classes of responses for opportunities of occurrence, we find that rewarded responses are recalled significantly more often than are punished ones ($t = 1.96$, $df = 179$, $P = .05$).²⁰ Our findings here differ from those of Wallach and Henle who found the level of recall for rewarded and punished responses almost identical. The general level of recall was, however, higher in our experiment than in that of Wallach and Henle. As in the case of repetitions, a high level of responding may be necessary to bring out the differential effectiveness of reward.

It is difficult to choose between several alternative interpretations of this finding, after the fact. In this type of experiment, Thorndike has not explicitly addressed himself to the problem of recall, but the general tenor of his analysis suggests that rewards should differentially strengthen both repetitions and recalls.²¹ Application of Tolman's principle of emphasis would also predict these results. Rewards, it has been asserted, are more effective emphasizees than punishments.²² Hence *S* acquires more information about correct than about wrong sequences of word-number during the original training. Finally, the Köhler-Restorff hypothesis concerning the effects of perceptual isolation on retention would also require the superior retention of the spaced infrequently rewarded responses.²³ The reader must choose among these alternatives according to his convictions.

Relation between recall and repetition. Performance-instructions significantly influence repetitions but not recall. The question remains to what degree repetitions and recalls are associated within conditions. Let us assume that most of the responses which the *S* can reproduce on the recall-test were also available to him during the test-series. Such an assumption is,

²⁰ There were no differences among conditions as regards the superior recall of rewarded responses (cf. Table IV).

²¹ As Thorndike recognized, the law of effect cannot strictly predict that rewarded responses will be remembered more frequently than punished responses. "The Law of Effect would not lead us to remember experiences that were pleasant and forget experiences that were painful, but to remember experiences that have been *pleasant to remember*, and to forget experiences that have been *painful to remember*, a very different matter" (*op. cit.*, 1932, 458 f.).

²² E. C. Tolman, C. S. Hall, and E. P. Bretnall, A disproof of the law of effect and the substitution of the laws of emphasis, motivation and disruption, *J. Exper. Psychol.*, 15, 1932, 601-614.

²³ H. von Restorff, Ueber die Wirkung von Bereichsbildung im Spurenfeld, *Psychol. Forsch.*, 18, 1933, 299-342.

to be sure, hazardous but receives some indirect support from the fact that there are no differences in recall between the experimental groups and the Recall-Control Group. The effects of performance-instructions should then produce the following correlations between number of repetitions and number of recalls. When performance-instructions are positive, we should expect a high degree of association between repetitions and recalls since

TABLE V
CORRELATIONS BETWEEN THE NUMBERS OF REPETITIONS AND RECALLS OF REWARDED
RESPONSES BY THE VARIOUS EXPERIMENTAL GROUPS

Punished responses	Instructions for					
	Rewarded responses					
	R_+		R_0		R_-	Totals
W_+	I .74	IV .73	VII .58			.69
W_0	II .56	V .59	VIII .43			.53
W_-	III .28	VI .28	IX .22			.26
Totals	.55	.56	.45			.52

S would be motivated to repeat all the responses he recalls. When performance-instructions are negative, the correlations should be low since S would be motivated not to repeat all the responses he recalls. By the same token, indeterminate instructions should produce an intermediate degree of correlation. Table V shows for the various conditions the correlations between repetition and recall of rewarded responses. Table VI presents the same information for punished responses. In averaging correlations for

TABLE VI
CORRELATIONS BETWEEN THE NUMBERS OF REPETITIONS AND RECALLS OF PUNISHED
RESPONSES BY THE VARIOUS EXPERIMENTAL GROUPS

Punished responses	Instructions for					
	Rewarded responses					
	R_+		R_0		R_-	Totals
W_+	I .86	IV .86	VII .74			.83
W_0	II .70	V .70	VIII .59			.66
W_-	III .60	IV .48	IX .19			.44
Totals	.74	.71	.54			.64

the different instructions (row and column averages), Fisher's z -transformation was used. All the correlations are positive, indicating that in general recall and repetition tend to be associated. There are, however, important differences among conditions.

Let us first consider the correlations between repetition and recall of rewarded responses. The performance-instructions for rewarded responses have only a small and insignificant effect on the size of these correlations

(column averages in Table V). The instructions for *punished* responses have a larger and significant influence on the correlation between repetitions and recalls for rewarded responses (row averages in Table V).²⁴ This finding is entirely consistent with the results presented in Table I. The pattern of correlations confirms our hypothesis of a general repetitional set determined by the instructions concerning the punished responses which form the majority of the series. The instructions for punished responses also significantly influence the correlation between repetition and recall for the punished responses themselves (row averages in Table VI). Instructions for rewarded responses do not significantly influence the correlations between repetition and recall of punished responses (column averages in Table VI). These findings are again in conformity with the hypothesis of a general repetition set.

Identification of reinforcements. On the recall-test *S* was required to state for each response whether it had been rewarded or punished during

TABLE VII
AVERAGE NUMBER OF RESPONSES IDENTIFIED AS REWARDED AND PUNISHED BY THE
EXPERIMENTAL GROUPS AND THE RECALL-CONTROL GROUP

Experimental Groups	Instruction	Rewarded responses called		Punished responses called	
		Right	Wrong	Right	Wrong
I	R ₊ W ₊	1.05	1.95	1.15	16.85
II	R ₊ W ₀	1.10	1.90	1.15	16.85
III	R ₊ W ₋	1.00	2.00	1.25	16.75
IV	R ₀ W ₊	.95	2.05	1.65	16.35
V	R ₀ W ₀	.70	2.30	1.50	16.50
VI	R ₀ W ₋	.75	2.25	1.20	16.80
VII	R ₋ W ₊	.55	2.45	1.30	16.70
VIII	R ₋ W ₀	.60	2.40	1.65	16.35
IX	R ₋ W ₋	1.05	1.95	1.35	16.65
All		.86	2.14	1.36	16.64
Recall-Control		.85	2.15	1.40	16.60

the training series. Table VII shows for all experimental conditions as well as for the Recall-Control Group the average numbers of correct and incorrect identifications of rewards and punishments. We note first of all that on the average 2.21 responses were called rewarded and 18.79 were called punished. Clearly, the *Ss* remembered that the large majority of responses had been punished. Hence their accuracy in identifying punished responses as punished was necessarily high: on the average 16.62 of the 18 punished responses were correctly identified. The identification of the

²⁴ In testing for significance of differences between correlations, the *z*-transformation was used.

rewarded responses was much less accurate. On the average, only 0.86 of the three rewards were correctly identified.

Even though accuracy in the identification of rewards is rather low, further analysis shows that the errors of identification are not entirely haphazard. The position in the original training series of each response called rewarded by *S* with respect to the response actually rewarded was determined. (It will be remembered that the rewards were so spaced in the series that each reward was preceded and followed by three punished responses.) When the average frequency with which responses were called 'Right' by *S* is plotted against distance from the rewarded response, the gradient shown in Fig. 1 results. The largest number of identifications is given for the actually rewarded response, and fre-

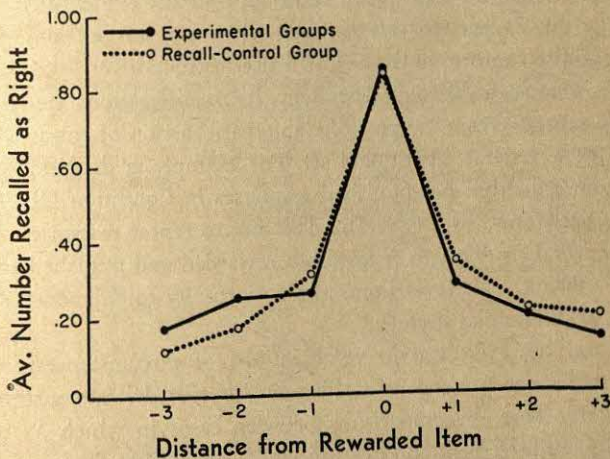


FIG. 1. AVERAGE NUMBER OF RESPONSES CALLED "RIGHT" AS A FUNCTION OF DISTANCE FROM REWARDED RESPONSE

quency of incorrect identifications varies inversely with the distance in the original training series from the rewarded position. This gradient of identification of rewards may be ascribed to the formation of remote associations between the reward and stimulus-words preceding or following the rewarded pair in the training series.²⁵ The greater the distance from the rewarded pair, the less likely are such remote associations to be formed. The difference between the rewarded positions and all other positions is highly significant ($F = 37.59$, $df = 2$ and 1074 , $P < .001$). The difference

²⁵ A 'spread of effect' in the recall of rewards was reported by J. Nuttin, 'Spread' in recalling failure and success, *J. Exper. Psychol.*, 39, 1949, 690-699.

between positions 1, 2, and 3 (combining those preceding and following the reward) are also significant at the 5-% level or better.²⁶

The differences between conditions in the identification of reinforcements are quite small and not significant. It is of particular interest to note that the average scores of the Recall-Control Group are almost identical with those of the combined experimental groups. Like recall, identification of rewards and punishments is independent of performance-instructions. These instructions are relevant to repetition only; they are no more relevant to the task of identifying rewards and punishments than they are to recall.

Evidence for strategy. So far the analysis of Ss' performance has produced evidence for the effectiveness of reward and for the operation of a general repetitional set. Can we, in addition, find any evidence for deliberate strategy on S's part to differentiate between rewarded and punished responses and to conform to the performance-instructions?

If there were such differential strategy, its manifestations are in general difficult to separate from the possible 'automatic' action of rewards and the operation of a general repetitional set. For example, a deliberate strategy to repeat rewarded but not punished responses in Condition III ($R_+ W_-$) may be 'canceled out' by a generalized set not to repeat responses. In addition, the general tendency to repeat both rewarded and punished responses at a better than chance level limits the opportunity for demonstrating the operation of differential strategy.

There are, however, certain combinations of circumstances in which critical tests for the presence of deliberate differential strategy seem possible. We refer here to comparisons between cases in which Ss correctly identify the reinforcement and cases in which they fail to do so.

Consider as an illustration two groups of rewarded responses reproduced correctly in the recall test by Ss in Condition III ($R_+ W_-$). One of these groups of responses is correctly identified as rewarded, whereas the other group is not, *i.e.* is identified as punished. As far as S's performance is determined only by the 'automatic' action of reward and a general repetition set, both groups of responses should be equally likely to be repeated. If, however, S pursues a deliberate differential strategy, he should repeat more of the responses identified as rewarded. Thus, evidence for or against differential strategy must be sought by an analysis of the interaction between recall, identification of reinforcement, and repetition.

²⁶ In testing the significance of differences between adjacent positions, we used the procedure developed by J. W. Tukey, Comparing individual means in the analysis of variance, *Biometrics*, 5, 1949, 99-114.

For purposes of this analysis we consider only responses that are correctly reproduced on the recall-tests. This limitation was imposed on the assumption that only responses recalled by *S* could be successfully used in the pursuit of a differential strategy.²⁷ For all such responses we then ask the questions whether correct identification of the reinforcement enhances conformity with instructions as compared to incorrect identification.

Such an analysis can be carried out only for those conditions in which performance-instructions for the class of responses in question are explicit and differ from those for the other class of responses. In the case of rewarded responses, Conditions II ($R_+ W_0$), III ($R_+ W_-$), VII ($R_- W_+$), VIII ($R_- W_0$) meet this criterion; in the case of punished responses, Conditions IV ($R_0 W_+$), VII ($R_- W_+$), and III ($R_+ W_-$) and VI ($R_0 W_-$) meet this criterion. Table VIII shows for rewarded responses that had

TABLE VIII
RELATION BETWEEN IDENTIFICATION OF REINFORCEMENT AND REPETITION
FOR REWARDED RESPONSES CORRECTLY RECALLED

Experimental Groups	Instruction	Called right by <i>S</i>		Called wrong by <i>S</i>	
		No. responses repeated	No. responses not repeated	No. responses repeated	No. responses not repeated
II	$R_+ W_0$	12	4	4	4
III	$R_+ W_-$	5	4	0	8
	Total R_+	17	8	4	12
VII	$R_- W_+$	4	5	12	5
VIII	$R_- W_0$	4	9	6	7
	Total R_-	8	14	18	12

been correctly reproduced in the recall-test the relationship between identification of the reinforcement by the *Ss* and repetition. Table IX presents the same information for punished responses. Since the numbers of cases which meet our criteria of selection are small, we present straight frequencies of occurrence of responses and do not attempt any statistical evaluation. In this case, our guesses concerning the significance of the results must rest on the internal consistency of the data.

Examination of Tables VIII and IX shows that in general conformity with the performance-instructions varies with the *Ss'* identification of the reinforcement. If performance-instructions for rewarded responses are positive, the rewarded responses which are correctly identified are repeated

²⁷ As before, we assumed, not without hesitation, that responses correctly reproduced on the recall-test were, at least in the majority of the cases, also available to the *Ss* during the test series.

relatively more often than those which are not correctly identified (Table VIII). When the performance-instructions for rewarded responses are negative, the rank order of repetitions is reverse. The relationship is less clear-cut in the case of punished responses since the great majority of the punished responses was correctly identified (Table IX).

The association between recall, identification of reinforcement, and repetition lends some support to the view that *S* attempts to pursue a deliberate differential strategy in carrying out his instructions. This conclusion is necessarily tentative since there was only a small number of responses from which

TABLE IX
RELATION BETWEEN IDENTIFICATION OF REINFORCEMENT AND REPETITION
FOR PUNISHED RESPONSES CORRECTLY RECALLED

Experimental Group	Instruction	Called Wrong by S		Called right by S	
		No. responses repeated	No. responses not repeated	No. responses repeated	No. responses not repeated
IV	W_+R_0	69	30	5	4
VII	W_+R_-	66	45	9	6
	Total W_+	135	75	14	10
III	W_-R_+	18	59	1	2
VI	W_-R_0	33	54	2	3
	Total W_-	51	113	3	5

unequivocal evidence for differential strategy could be obtained. In any event, such strategy could not possibly account for all the variations in *S*'s performance in view of the positive evidence for the operation of a general repetitional set and the 'automatic' action of reward.

DISCUSSION

Two points are clear: (1) the *Ss* do learn both rewarded and punished word-number sequences in the absence of instructions to learn; (2) the performance-instructions given before the test significantly influence performance. These findings confirm the importance of the conceptual distinction between learning and performance and demonstrate the relevance of this distinction to the rôle of reinforcement in verbal behavior. At the same time, our results point to important limitations in the application of this distinction.

In the analysis of the principles governing learning and performance two assumptions are frequently made: (1) that rewards and punishments function only as perceptual emphasizees and do not establish dispositions

to respond in specific ways,²⁸ and (2) that once an organism has learned he will use his knowledge 'sensibly' in accordance with the demands of the situation.²⁹ These assumptions may be correct in some situations but it seems to us that they exaggerate the differences between the conditions of learning and performance. We have clear evidence in our results that rewarded responses are repeated at a better than chance level even though performance instructions do not call for such repetitions. In fact, we find, as Thorndike did, that punished responses too are repeated at a better than chance level, as if by virtue of their sheer occurrence.³⁰ Rewards are, however, in general more effective than punishments. The general picture that emerges is that of performance-instructions modulating, over a wide range, what appear to be automatic effects of reward and exercise.

As we have pointed out repeatedly, the effects of instructions cannot be fully understood in terms of a deliberate differential strategy pursued by the *S*. While there is some tentative evidence for such strategy, the single most important performance-effect is the operation of general repetitional sets. Instructions concerning the relatively more numerous punished responses influence the level of repetition of both rewarded and punished responses. A logical analysis of what the sensible course of action is for *S* surely would not have predicted this result. Such principles of performance must, it appears, wait for empirical discovery.

The question has often been raised whether the law of effect is a law of learning or a law of performance. On the basis of our analysis we must conclude that it is both. This conclusion is dictated by the experimental facts thus far obtained; it may have to be revised when the principles of performance are further understood.

SUMMARY

The distinction between learning and performance is of critical importance for the theoretical analysis of the effects of reinforcement. The experiment reported in this paper applies this distinction to a situation which has often been used to test Thorndike's law of effect—the learning of word-number associations under differential reinforcement. The experiment was designed to answer two questions: (1) Does learning take place even though *S* is not motivated to learn, *i.e.* under conditions which make the rewards and punishments irrelevant to *S*'s task? (2) If learning does take place under such conditions, do different instructions concerning repetition

²⁸ Tolman, Hall, and Bretnall, *op. cit.*, 609-614.

²⁹ Wallach and Henle, *op. cit.*, 1941, 348 f.

³⁰ Thorndike, *op. cit.*, 1932, ch. 11.

of rewarded and punished responses produce significant variations in performance?

A series of 25 three-letter words was presented to our Ss with instructions to respond with a number from 1 to 10. There were a training series and a test-series. Prior to the training series the Ss were informed that the experiment tested success in extra-sensory perception and that the numbers assigned to the words would vary at random from series to series. Responses in positions 6, 13, and 20 were arbitrarily called 'Right' during the training series, all other responses were called 'Wrong.' The list of words was then repeated for the test-series, but in a different order. Prior to the test-series, the experimental groups were given different instructions designed to manipulate the Ss' motivation to repeat or not to repeat responses rewarded or punished during the test-series. Different groups were informed that, for the test-series, rewarded or punished responses, or both, would (a) be correct, (b) be incorrect, or (c) might or might not be correct. Each of nine experimental groups received one of the possible combinations of instructions. Following the test-series, a recall-test was given on which the Ss recalled the responses they had given during the training series. The Ss were also required to recall whether their responses had been rewarded or punished. A Recall-Control Group was given a recall-test immediately after the training series.

Analysis of the Ss' repetitions and recalls of the numbers rewarded and punished during the training series yielded the following conclusions.

(1) Both rewarded and punished responses are repeated at a better than chance level under most conditions, even when performance-instructions do not call for such repetition.

(2) The level of repetition is significantly influenced by performance-instructions. The critical instructions are those for punished responses which refer to the majority of items in the list. The more positive the instructions for punished responses, the higher the level of repetition of *both rewarded and punished responses*. These effects of performance-instructions are interpreted as evidence for a general repetition set.

(3) Rewarded responses are repeated relatively more often than are punished responses. This difference is significant if and only if there is no instruction against repetition of rewarded responses and if the instructions for punished responses are not negative. The differential effectiveness of rewards, in conjunction with the better than chance repetition of rewarded responses under most conditions, is in agreement with Thorndike's hypothesis of the 'automatic' action of reinforcement.

(4) There is no difference among the experimental conditions in the number of responses correctly recalled, nor are the experimental groups different from the Recall-Control Group. Performance instructions influence frequency of repetitions but not recall.

(5) Rewarded responses are recalled relatively more frequently than punished responses.

(6) There is in general a positive correlation between the number of responses repeated and recalled. The size of the correlation varies, however, with performance-instructions. The stronger the repetitional set (induced by instructions for punished responses), the higher is the correlation between repetition and recall.

(7) A smaller percentage of rewards than punishments is correctly identified. There is no significant difference among the experimental groups in the accuracy with which reinforcements are identified, nor do the experimental groups differ from the Recall-Control Group in the accuracy of identification.

(8) There is a gradient in the identification of rewards. The largest number of identifications occurs for the actually rewarded response. The frequency of incorrect identifications varies inversely with the distance from the rewarded response.

(9) There is some evidence for a deliberate strategy on the part of the Ss to differentiate between rewarded and punished responses and to conform with the performance-instructions. The effectiveness of such a strategy is necessarily limited by the 'automatic' action of rewards and the influence of general repetition sets.

The most general conclusion from the results of the experiment is that performance-instructions modulate, over a wide range, the effects of reinforcement and exercise. The distinction between learning and performance is clearly relevant to the analysis of the rôle of reinforcement in verbal behavior.

THE EFFECT ON CRITICAL FLICKER FREQUENCY OF INTEROCULAR DIFFERENCES IN INTENSITY AND IN PHASE RELATIONS OF FLASHES OF LIGHT

By GARTH J. THOMAS, University of Chicago

In a classic experiment published in 1904, Sherrington investigated the effect on critical flicker frequency (*CFF*) of various phase relations of the flashes of light presented to the two eyes and the effects of interocular differences in intensity.¹ He was interested in determining whether the nervous system followed the same laws of integration on the sensory side as his investigations of spinal reflexes had indicated for the motor side and as was implied by Cajal's theory.² Cajal had postulated what he called 'isodynamic' cells in the cortex on which excitation from corresponding points in the two retinas converged. The function of these isodynamic cells represented the basis of unified binocular vision. Sherrington found the effects of the experimental variables in his study on *CFF* to be very slight and insignificant, and he concluded, contrary to Cajal's theory, that binocular perception results from the combination of "already elaborated unocular sensations contemporaneously proceeding."³ He suggested that the phenomena of binocular fusion take place *psychically* and are not based on a physiological mechanism analogous to the integrative pool of neurons at the entrance of final common paths in the motor system.

The purpose of the present study is to determine the effects on binocular *CFF* of differences of illumination on corresponding areas of the two eyes and a comparison of in- and out-of-phase flashes, with stimulus-lights of two sizes and each located centrally and peripherally on the retinas.

* Accepted for publication September 15, 1953. This work was supported in part by the Medical Research and Development Board, Office of the Surgeon General, Department of the Army, Contract No. DA-49-007-MD-36. In the early stages the study was supported by a grant from the Dr. Wallace C. and Clara Abbott Memorial Fund of the University of Chicago. I wish to express my appreciation for the assistance of Mr. Wallace Welker and Mr. Theodore Schaefer.

¹ C. S. Sherrington, On binocular flicker and the correlation of activity of 'corresponding' retinal points, *Brit. J. Psychol.*, 1, 1904, 26-59.

² Santiago Ramón y Cajal, *Histologie du Système Nerveux de l'Homme et des Vertébrés*, 2, 1911, 880.

³ Sherrington, *The Integrative Action of the Nervous System*, 1906, 354-386.

Sherrington investigated the effects of differences in interocular phase of flash and the effects of a steady light in one eye on *CFF*, but he did not determine these interocular effects over wide ranges of stimulus-intensity at retinal locations other than foveal, or with stimuli of various angular subtense. He reported critical rates as high as 90-120 flashes per sec. In terms of subsequent experimental evidence regarding maximal critical rates, these figures are too high (almost by a factor of two), suggesting that considerable instrumental error must have been involved.⁴ In their investigation of Fechner's paradox, De Silva and Bartley found that processes from the more dimly illuminated eye seem to subtract from the magnitude of the process in the more highly illuminated eye when they are integrated in the central cyclopean retina.⁵ If the cyclopean image is a fusion of already elaborated unocular sensations proceeding independently, as Sherrington concluded, and if brightness is a function of intensity independent from flicker discrimination, as Crozier and Wolf have argued, then the dimmer half-image should have no effect on the brighter half-image as regards *CFF*.⁶

The different sized targets at central and peripheral retinal locations were introduced to tap differential proportions of the duplex receptor-system. One might surmise from present knowledge that in the function of the binocular system, foveal cones and their associated central processes play a more important rôle in binocular integration than does the more diffuse peripheral system. The fovea-periphery gradient of the size of Panum's fusional areas,⁷ and the chronic nystagmus and poor binocular fusion in some types of total color blindness in which functional cones are apparently lacking suggest such a notion.⁸

The third variable, phase of flash, was introduced to determine how its effects would manifest themselves as the two half-images approach equality in intensity. Several recent studies have shown that synchronous flashes of equal intensity in the two eyes yield higher *CFF* than that obtained with alternate flashes or with unocular regard.⁹ One would not expect phase of flash to have any effect when the retinal illuminance in the two eyes is widely different because the flash-rate of the less intensely illuminated eye would be well above its *CFF*. If the locus of binocular interaction of bursts of impulses occasioned by the flashes of light is earlier in the chain of cortical events than the locus of flicker fusion, one might expect phase effects to appear when one light is still below its *CFF*.

⁴ S. H. Bartley, *Vision: A Study of Its Basis*, 1941, 122-123.

⁵ H. R. De Silva and S. H. Bartley, Summation and subtraction of brightness in binocular perception, *Brit. J. Psychol.*, 24, 1930, 538-541.

⁶ W. J. Crozier and Ernst Wolf, Theory and measurement of visual mechanisms: IV. Critical intensities for visual flicker, monocular and binocular, *J. Gen. Physiol.*, 24, 1941, 505-535.

⁷ K. N. Ogle, *Research in Binocular Vision*, 1950, 64-68.

⁸ D. B. Judd, Current views on colour blindness, in F. P. Fischer, A. J. Schaeffer, and A. Sorsby (eds.), *Documenta Ophthalmologica: Advances in Ophthalmology*, 3, 1949, 257-258.

⁹ F. H. Ireland, A comparison of critical flicker frequencies under conditions of monocular and binocular stimulation, *J. Exper. Psychol.*, 40, 1950, 282-286; C. H. Baker, The dependence of binocular fusion on timing of peripheral stimuli and on central processes: 1. Symmetrical flicker, 2. Asymmetrical flicker, 3. Critical flicker, *Canad. J. Psychol.*, 6, 1952, 1-10, 84-91, 123-130; and 151-163; and F. H. Perrin, A study in binocular flicker, *J. Opt. Soc. Amer.*, 44, 1954, 60-69.

METHOD AND PROCEDURE

Apparatus. Fig. 1 shows a schematic diagram of the apparatus. Two front surface mirrors (*FSM*) reflect light from a 300-w. projection-bulb (*PB*) into parallel tubes containing lenses (the *Ls*) which focus images of the filament in the plane of the sector disks (*SD*). The power supply for the light source was a direct current generator. As the light was the only load on the *DC* line and the generator was driven by a synchronous motor, the power supply was very stable. The bulb was always operated at 2.5 amp. After their points of focus in the plane of the sector disks, the diverging light beams then pass through neutral tint optical wedges and balancers and filters (*W*, *BW*, and *F*) and form homogeneous spots of light on two milk glass surfaces (*MG*). Metal templates with apertures determining the size, shape, and location of the stimuli in the visual field were inserted in front of the milk glass. *S* views the two circular patches of light, one with each

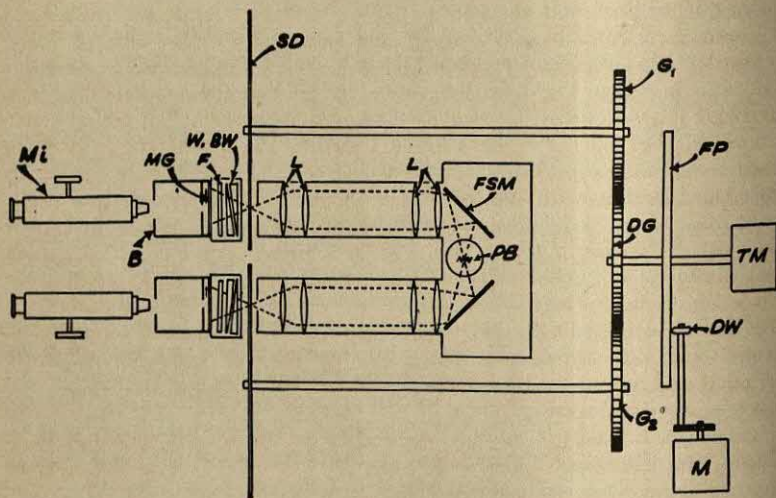


FIG. 1. DIAGRAM OF THE APPARATUS

eyes, through two parallel, low powered microscopes (*Mi*) which are focused on the apertures in the templates and whose horizontal separation was adjustable for interpupillary distance. Between the object lens of each microscope and its target (the metal template containing the apertures) was mounted a small box, lined with black flock (*B*) and containing a microscopic cover slip placed at an angle of 45° to the line of sight, which served to reflect a tiny red point into each field of view for fixation. The fixation-points were adjustable in intensity and could be located anywhere in their respective fields.

The sector disks rotate on axles connected by gears (*G*) to a central driving gear (*DG*). To the axle of this driving gear was fastened a metal friction plate (*FP*). A small leather drive wheel (*DW*), driven by a constant speed synchronous motor (*M*), pressed against the friction plate, causing it to rotate. Continuous adjust-

ment of the rate of rotation was accomplished by moving the motor and the drive wheel assembly along a horizontal radius of the friction plate by means of a screw adjustment. Phase relations in the delivery of flashes to the two eyes were determined by the angular displacement on their respective axles of the sector disks relative to each other. Only two phase relations were used, synchronous, in-phase flashes and alternate, out-of-phase flashes. The rate of rotation was measured by means of a Weston Tachometer (TM) which had been carefully calibrated against a Strob-o-tac. The wedges and filters were calibrated in the apparatus with a Macbeth Illuminometer.

Subjects. Two Ss (TS and GT) served in the experiments. Both were highly practiced in the observations.

Procedure. Essentially the procedure consisted of determinations of CFF under conditions of binocular and uniocular regard and with variable differences in illumination on corresponding areas of the two eyes. CFF was determined by a modified method of limits in which the flash rate was set well above fusion and then reduced by small steps, delivered at S's signal, until he indicated he had detected flicker. S signaled by pressing microswitches held in his hands which turned on the lights on the control panel. The movements required to press the switches were slight and, as S's elbows were braced on a support, the signaling response did not noticeably disturb the orientation of his head and eyes. S's head was held in position by biting on a dental wax impression of his teeth and by a support pressing against his forehead. Great care was taken to align S's head that beams from the exit pupils of the microscopes were centered in the pupils of his eyes. Thresholds were taken with circular stimuli of 6° and 2.2° angular subtense, both with central fixation and viewed in the periphery. For peripheral stimulation the targets appeared in the lower left quadrant of each visual field, 10° from their centers to the fixation-points. The two stimuli and their associated fixation-points were located on corresponding retinal areas and therefore were seen as fused in binocular regard. With each size of target and retinal location thresholds were taken with flashes in phase and out of phase in the two eyes.

For each condition the stimulus in the right eye was fixed in luminance at an intermediate level on the log I scale, and the flicker threshold of that eye was determined from time to time during the course of a day's run. The light in the left eye was of variable intensity and at each level flicker thresholds were measured with both uniocular and binocular regard. The light-to-dark ratio was one. Before each experimental session, S was dark adapted. Settings of the level of luminance of the variable (left) light always progressed from lower to higher, and a short period of time for light adaptation was allowed at each new level. CFFs of the right eye alone at an intermediate level of luminance, of the left eye alone at various levels of luminance over a range of about 6 log units, and with both eyes at each level of luminance of the left light were all determined at one session.

RESULTS AND DISCUSSION

Fig. 2 and Table I show the results for both Ss with the 6° target centrally fixated. The values on the ordinate are flashes per second and on the abscissa, flash luminance in log millilamberts. The solid circles represent the CFF contour for the left eye alone. The solid line was fitted by inspection.

tion. The square near the center of the curve represents the log luminance of the stimulus and the corresponding *CFF* for the right eye alone. The open circles connected by dashed lines represent *CFF* with binocular regard, and are plotted in terms of the luminance of the left stimulus. The binocular thresholds on the two upper graphs were taken with in-phase flashes in the two eyes, and on the two lower ones with out-of-phase flashes. Every point on all graphs represents the mean of five determinations. Evidence for binocular interaction is shown by the discrepancy between

TABLE I

CRITICAL FLICKER FREQUENCY AS A FUNCTION OF LOG LUMINANCE FOR LEFT AND RIGHT EYES ALONE AND FOR BOTH EYES WITH IN- AND OUT-OF-PHASE FLASHES (6° target, central fixation)

Results of TS					Results of GT				
Log I_{mL}	In phase		Out of phase		Log I_{mL}	In phase		Out of phase	
	left	both	left	both		left	both	left	both
-2.160	11.3	30.0	10.0	28.5	-2.176	10.1	31.2	10.2	30.8
-1.714	11.4	30.2	10.2	28.6	-1.734	11.4	31.4	12.0	30.7
-1.322	12.9	29.9	12.3	28.3	-1.320	12.4	31.7	12.0	30.2
-0.917	13.2	29.7	13.1	28.8	-0.947	12.9	31.3	14.2	31.1
-0.581	15.5	29.6	14.7	28.5	-0.608	16.1	31.6	15.2	30.9
-0.016	22.4	30.6	21.3	29.3	-0.025	24.3	32.4	21.6	31.2
0.430	27.5	31.2	25.5	29.8	0.417	30.2	32.6	28.7	30.9
0.726	31.2	33.0	30.2	29.4	0.643	32.0	35.0	31.8	31.4
0.823	32.4	33.9	31.7	30.5	0.833	34.1	34.8	32.4	31.2
1.227	36.6	36.1	36.0	33.9	1.204	38.3	36.8	35.4	34.2
1.563	41.0	40.1	39.8	37.8	1.543	40.8	39.5	39.5	37.2
2.086	46.1	44.9	43.6	43.7	2.069	45.9	42.9	43.8	43.2
2.532	51.2	49.1	49.2	47.4	2.512	49.5	47.4	47.9	46.7
2.923	56.0	53.7	53.4	51.2	2.927	52.1	49.7	50.7	50.3
3.329	56.7	54.7	55.4	52.7	3.298	51.8	49.5	51.9	49.2
3.665	56.6	54.2	54.2	51.2	3.638	49.5	47.9	53.7	50.7

Right eye alone:

Log I_{mL} = 0.758
CFF (in-phase) = 31.3
CFF (out-of-phase) = 30.0

Right eye alone:

Log I_{mL} (in-phase) = 0.674
CFF (in-phase) = 33.4
 Log I_{mL} (out-of-phase) = 0.776
CFF (out-of-phase) = 32.1

binocular thresholds and the thresholds for the more intensely illuminated eye alone. If the processes from homonymous hemiretinas do in fact function independently with regard to the discrimination of flicker, binocular thresholds would be the same as the uniocular thresholds of the more intensely illuminated eye. For about three log units on the abscissa from the left, the right stimulus was the more intense. If for binocular regard, the critical flash rate were indeed determined by the more intensely illuminated eye, the points should cluster around the horizontal line running from the square to the ordinate. To the right of the square, the left stimu-

lus was more intense than the right and there should be no consistent difference between the uniocular (solid circles) and the binocular (open circles) thresholds. Instead, there is a consistent subtractive or damping effect of the more dimly illuminated eye on the brighter as regards the detection of flicker using both eyes, *i.e.* a steady dim light in one eye decreases the sensitivity of the corresponding area of the other eye for detection of flicker. Phase of flash has a differential effect only when the

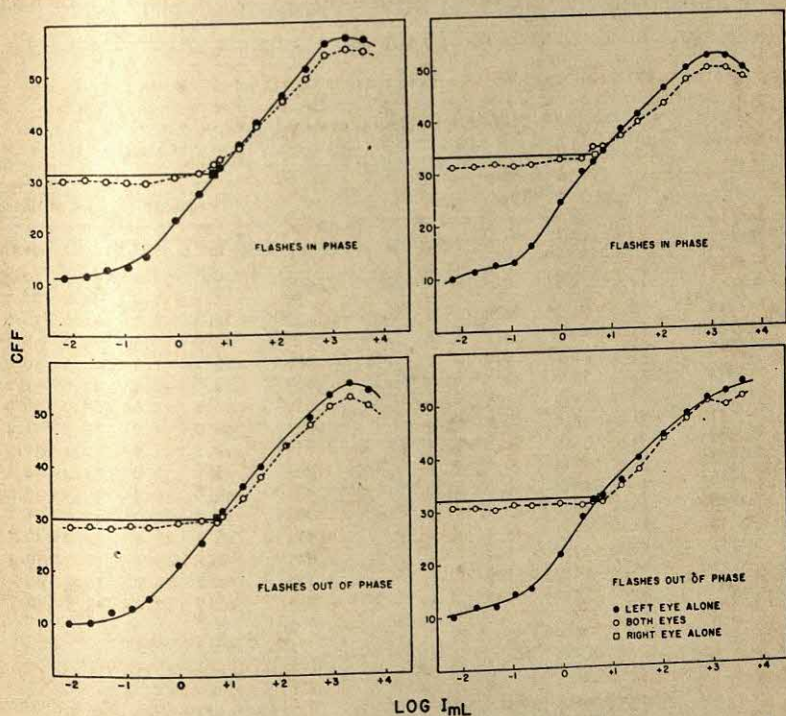


FIG. 2. CFF AS A FUNCTION OF LOG LUMINANCE
(6° target; central fixation; results of TS to the left, of GT to the right)

intensities of the two half-images approach equality. Flashes in phase in the two eyes yield higher CFF than do flashes out of phase.

Fig. 3 and Table II show similar results for the smaller stimuli (2.2° subtense), centrally fixated. The magnitude of the subtractive effects appears, however, to be less, especially for values of lower luminance. For both Ss, when the left light is dimmer than the right, binocular thresholds are essentially the same as the uniocular threshold of the more intensely illuminated eye. In the upper halves of the curves, where the left light is

more intense than the right, there is a consistent small depression of *CFF* with both eyes as compared with *CFF* of the left eye alone.

With 6° targets located 10° peripherally from the fovea, the subtractive effect is more clear-cut but more irregular as shown in Fig. 4 and Table III. Results for the 2.2° stimulus in the periphery (Fig. 6, Table IV) are similar to those with the small stimuli viewed in central fixation in that at lower intensities there is no consistent subtractive effect, but there is a more marked effect at higher intensities. The last few points which start to rise again on the curves for *TS* are most certainly artifacts due either

TABLE II

CRITICAL FLICKER FREQUENCY AS A FUNCTION OF LOG LUMINANCE FOR LEFT AND RIGHT EYES ALONE AND FOR BOTH EYES WITH IN- AND OUT-OF-PHASE FLASHES (2.2° target, central fixation)

Log I_{mL}	Results of <i>TS</i>				Results of <i>GT</i>			
	In phase		Out of phase		In phase		Out of phase	
	left	both	left	both	left	both	left	both
-1.921	—	—	—	—	5.0	29.6	—	—
-1.734	8.0	27.5	8.6	27.4	8.6	30.0	5.5	28.6
-1.320	10.1	27.6	10.3	27.7	10.8	29.7	9.3	29.5
-0.947	11.9	27.4	12.2	27.8	13.2	29.6	11.9	29.9
-0.608	14.9	27.4	15.0	27.6	15.2	28.5	16.7	30.5
-0.025	19.4	26.6	20.1	26.4	20.7	30.0	21.5	28.8
0.417	24.5	28.1	24.2	26.4	25.7	30.5	27.5	28.7
0.691	27.8	29.3	27.8	26.4	30.0	30.9	29.9	28.5
0.833	29.0	29.6	29.2	28.0	30.8	32.6	31.7	30.2
1.204	32.1	31.7	31.4	30.2	35.3	33.8	34.1	32.9
1.543	35.4	34.2	34.8	32.9	37.5	36.5	37.7	35.7
2.069	39.6	38.3	39.3	38.6	42.0	40.4	42.0	39.9
2.512	42.8	40.7	42.2	41.0	46.2	44.7	46.7	44.6
2.927	45.0	43.7	44.3	41.9	48.6	46.7	49.7	48.0
3.298	46.2	44.6	42.3	41.4	50.1	49.8	52.8	50.6
3.638	45.0	42.9	43.2	41.9	51.0	48.3	51.0	49.4

Right eye alone:

Log I_{mL} = 0.673

CFF (in-phase) = 27.9

CFF (out-of-phase) = 27.7

Right eye alone:

Log I_{mL} = 0.673

CFF (in-phase) = 30.1

CFF (out-of-phase) = 30.4

to failure to maintain fixation which was very difficult and very critical, especially at higher intensities, or to the spread of stray light, both entopic and within the instruments, to areas of the retina more sensitive to flicker, thus causing the whole field to appear to flicker.

Figs. 6 and 7 show all the results plotted in terms of the difference between binocular *CFF* and that of the more intensely illuminated eye as a function of the difference in luminance of the right and left stimuli. The abscissas for all graphs is the difference in log units (not the log of the differences) between the luminance of the stimuli in the two eyes. The

negative signs simply indicate that the right eye is the more intensely illuminated one. The zero in the middle is the point of no difference, or equality, of luminance in the two eyes, and the positive values indicate the left eye is the more brightly illuminated. The ordinates of all curves represent the difference in flashes per second between thresholds with binocular regard and the uniocular thresholds of the more intensely illuminated eye. Positive values indicate summation, *i.e.* binocular thresholds higher than the uniocular threshold of the more intensely illu-

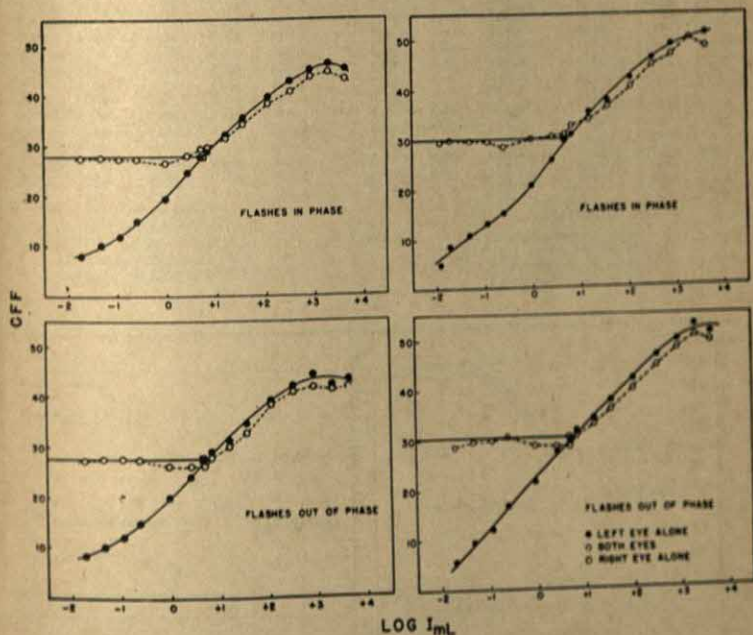


FIG. 3. CFF AS A FUNCTION OF LOG LUMINANCE

(2.2° target; central fixation; TS's results to the left, GT's to the right)

minated eye; negative values indicate a subtractive or damping effect in which binocular thresholds are lower than the uniocular thresholds of the more intensely illuminated eye. Open circles represent flashes out of phase; solid circles, flashes in phase. The eight individual graphs represent the results of the four conditions of size of target and retinal location for both Ss.

The most striking fact shown by these graphs is the increased sensitivity for detection of flicker with flashes in phase in the two eyes as the luminance of the two half-images approaches equality. In every case thresholds

TABLE III

CRITICAL FLICKER FREQUENCY AS A FUNCTION OF LOG LUMINANCE FOR LEFT AND RIGHT EYES ALONE AND FOR BOTH EYES WITH IN- AND OUT-OF-PHASE FLASHES (6° target, 10° peripheral fixation)

Results of TS						Results of GT					
In phase			Out of phase			In phase			Out of phase		
Log $I_m L$	left	both	Log $I_m L$	left	both	Log $I_m L$	left	both	Log $I_m L$	left	both
-2.450	15.1	27.8	-2.477	16.1	28.4	-2.477	14.4	28.9	-2.477	15.2	24.8
-2.176	16.9	27.7	-2.204	16.7	26.1	-2.204	15.3	28.9	-2.204	15.2	25.5
-1.734	17.6	24.4	-1.762	18.0	26.3	-1.762	15.5	27.7	-1.762	16.4	27.2
-1.320	18.9	24.8	-1.347	18.0	25.5	-1.347	16.9	27.6	-1.347	16.7	26.9
-0.947	18.0	25.5	-0.974	17.9	26.5	-0.974	16.0	26.0	-0.974	16.5	26.0
-0.608	18.0	25.6	-0.635	17.3	26.1	-0.735	16.3	26.0	-0.735	17.0	25.7
-0.025	19.8	26.6	-0.052	19.2	26.3	-0.052	18.8	26.0	-0.052	19.8	26.6
0.417	27.1	29.1	0.390	24.8	27.5	0.390	24.0	28.8	0.390	22.8	25.5
0.691	29.3	30.8	0.664	28.5	28.5	0.664	27.3	29.6	0.664	28.7	28.1
0.805	32.4	32.4	0.805	30.3	28.5	0.805	29.4	30.5	0.805	30.2	30.3
1.176	34.8	35.0	1.176	35.0	32.6	1.176	32.7	32.3	1.176	35.0	33.8
1.516	41.6	38.9	1.516	37.8	37.2	1.516	39.0	36.0	1.516	36.8	36.2
2.042	44.7	44.1	2.042	43.2	41.3	2.042	41.4	38.7	2.042	44.1	41.3
2.484	47.7	44.7	2.484	46.1	44.4	2.484	45.0	36.5	2.484	46.8	43.8
2.900	47.4	43.8	2.900	46.2	43.2	2.900	44.9	37.1	2.900	46.5	42.3
3.271	47.0	44.6	3.271	42.9	42.2	3.271	42.0	38.4	3.271	44.1	39.0
3.610	46.7	45.9	3.610	44.7	42.9	3.610	44.1	36.8	3.610	46.2	40.2

Right eye alone:

Log $I_m L$ (in-phase) = 0.592

CFF (in-phase) = 28.1

Log $I_m L$ (out-of-phase) = 0.674

CFF (out-of-phase) = 28.2

Right eye alone:

Log $I_m L$ = 0.694

CFF (in-phase) = 28.6

CFF (out-of-phase) = 27.6

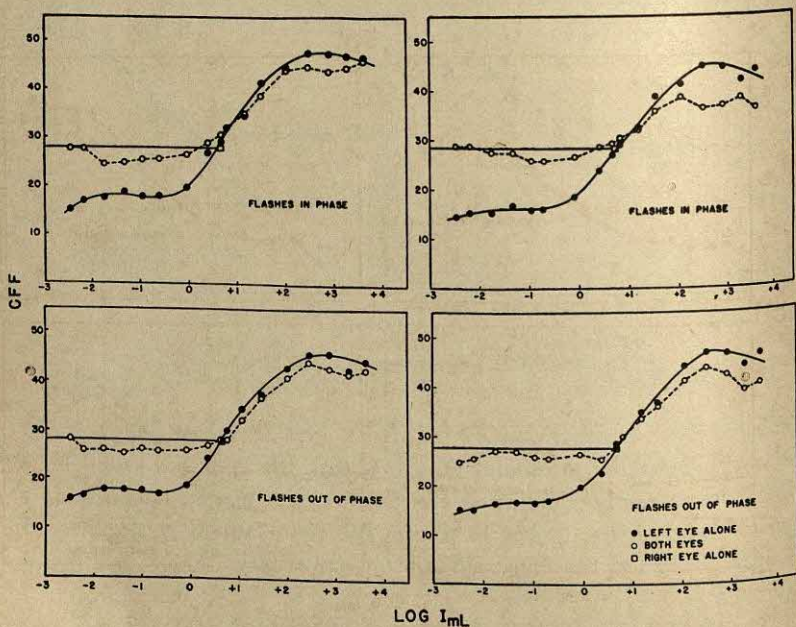


FIG. 4. CFF AS A FUNCTION OF LOG LUMINANCE (6° target; 10° peripheral fixation; TS's results to the left, GT's to the right)

TABLE IV

CRITICAL FLICKER FREQUENCY AS A FUNCTION OF LOG LUMINANCE FOR LEFT AND RIGHT EYES ALONE AND FOR BOTH EYES WITH IN- AND OUT-OF-PHASE FLASHES (2.2° target, 10° peripheral fixation)

Results of TS					Results of GT				
Log I_{mL}	In phase		Out of phase		Log I_{mL}	In phase		Out of phase	
	left	both	left	both		left	both	Log I_{mL}	left both
-2.450	11.2	19.7	9.6	18.3	-2.477	10.3	18.8	-2.450	10.2 19.2
-2.176	10.8	18.2	10.1	18.3	-2.204	11.0	20.7	-2.176	11.5 18.5
-1.734	10.7	17.9	10.9	17.8	-1.762	12.1	18.3	-1.734	11.9 18.5
-1.320	12.4	17.8	11.7	17.7	-1.347	12.6	17.8	-1.320	12.6 20.3
-0.947	12.1	18.3	12.6	18.1	-0.974	11.8	18.4	-0.947	11.5 20.2
-0.608	12.8	18.7	13.0	17.9	-0.635	12.7	17.5	-0.608	12.5 19.9
-0.025	15.6	18.6	16.1	17.6	-0.052	15.2	19.3	-0.025	13.7 17.3
0.417	18.4	20.1	18.5	17.9	0.390	17.3	20.4	0.417	16.8 17.9
0.691	19.9	20.9	20.6	19.4	0.664	19.6	21.8	0.691	19.4 18.2
0.833	21.3	22.7	21.9	20.6	0.805	20.5	22.2	0.833	20.4 20.4
1.204	24.0	23.6	24.5	23.3	1.176	24.3	24.2	1.204	26.1 24.5
1.543	29.6	27.8	26.6	24.3	1.516	28.4	26.1	1.543	28.4 27.2
2.069	31.7	28.1	27.6	24.9	2.042	31.5	29.7	2.069	33.0 30.6
2.512	32.6	27.9	27.6	24.8	2.484	37.5	32.7	2.512	38.6 35.3
2.927	30.8	29.7	28.7	27.3	2.900	39.8	36.8	2.927	40.2 34.1
3.298	33.2	30.9	30.3	28.5	3.271	39.8	36.3	3.298	36.5 32.0
3.638	34.1	32.0	34.4	31.7	3.610	37.7	35.6	3.638	40.1 36.6

Right eye alone:

Log I_{mL} = 0.410

CFF (in-phase) = 18.2

CFF (out-of-phase) = 18.2

Right eye alone:

Log I_{mL} (in-phase) = 0.694

CFF (in-phase) = 19.9

Log I_{mL} (out-of-phase) = 0.675

CFF (out-of-phase) = 19.7

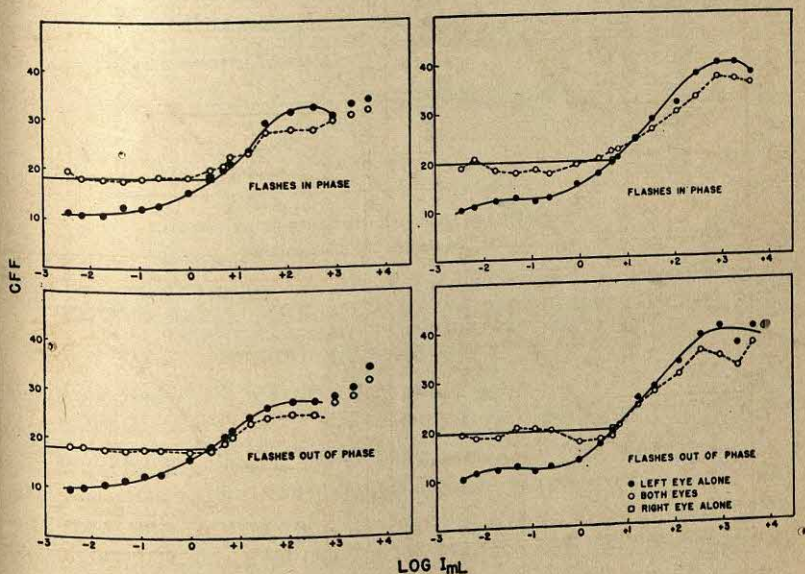


FIG. 5. CFF AS A FUNCTION OF LOG LUMINANCE (2.2° target; 10° peripheral fixation; TS's results to the left, GT's to the right)

determined with approximately equal luminance and synchronous flashes in the two eyes are higher than either of the corresponding uniocular thresholds or binocular thresholds with alternate flashes. Binocular thresholds with out-of-phase flashes tend to be the same as the uniocular thresholds, or in some instances, lower than either. For the four conditions of size of target and retinal location, in-phase flashes of approximately equal luminance in the two eyes yield an 8–12% increase over the out-of-phase

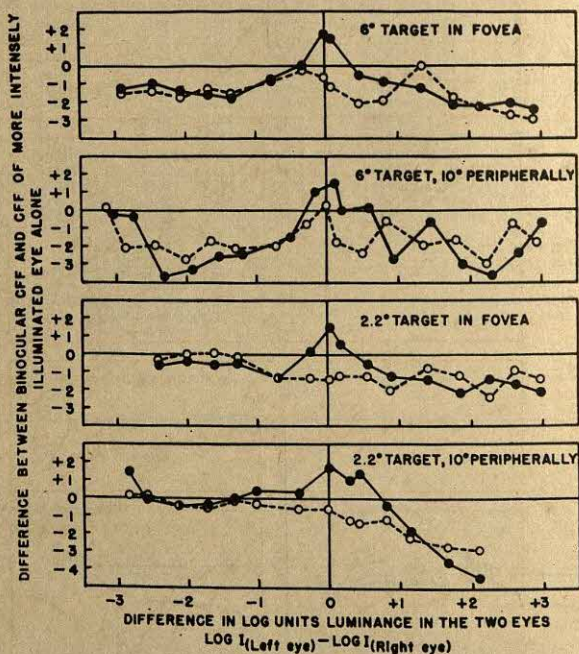


FIG. 6. DIFFERENCES BETWEEN BINOCULAR CFFs AND THOSE OBTAINED WITH THE MORE INTENSELY ILLUMINATED EYE AS A FUNCTION OF THE DIFFERENCE IN LUMINANCE BETWEEN THE RIGHT AND LEFT STIMULI
(Results of *TS*)

thresholds for *TS*. The corresponding increments for *GT*, whose thresholds are in general more erratic, are 5–20%.

In all instances the points for in-phase flashes indicate beginning interactions between excitations from homonymous hemiretinas by gradually rising over sometimes as much as a difference of a log unit in luminance of the half-images, to the peak at approximately equal luminance. This fact would seem to indicate that synchrony of flashes in the two eyes begins to have an effect while the flash rate is still well above fusion for the less intensely illuminated eye. That is, even though the dimmer light appears steady, its

intermittency must nonetheless be effective in arousing periodic bursts of activity in the homonymous nervous pathway which summates with the synchronous periodic excitation elicited by the more intense flash in the other eye. This interpretation is not, however, secure because the gradual rise of the in-phase curves is confounded with the decreasing subtractive effect attributable to the decreasing interocular differences in luminance.

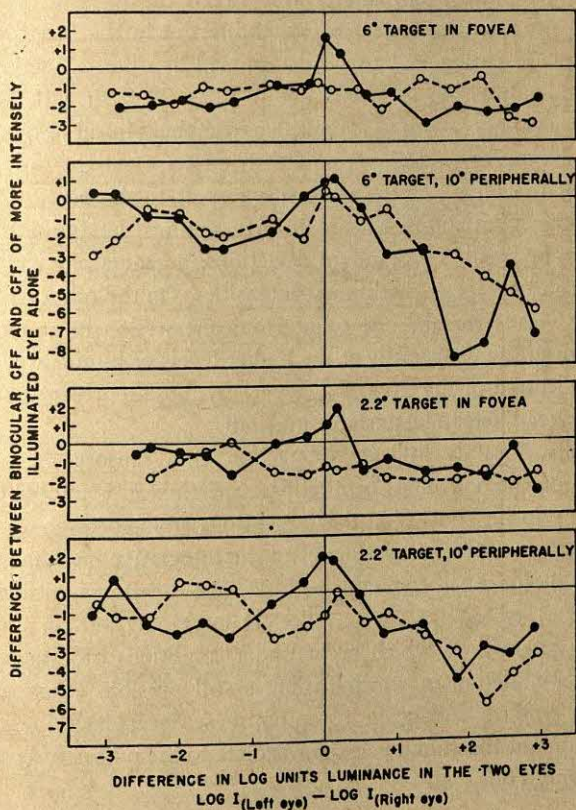


FIG. 7. DIFFERENCES BETWEEN BINOCULAR CFFs AND THOSE OBTAINED WITH THE MORE INTENSELY ILLUMINATED EYE AS A FUNCTION OF THE DIFFERENCE IN LUMINANCE BETWEEN THE RIGHT AND LEFT STIMULI (Results of GT)

Nonetheless, for about a quarter of a log unit difference on either side of equality, the in-phase binocular threshold tends to be higher than the CFF of the more intensely illuminated eye alone (zero on the ordinates), suggesting that the synchronous excitation can interact earlier in the chain of neural events than the neural locus of flicker fusion.

In all conditions except those with the smaller stimuli at lower intensities, differential luminance of the two half-images tends to yield lower *CFF* compared to those determined by the more intensely illuminated eye alone. The magnitude of the damping effect does not seem to be consistently related to the amount of the difference in luminance in the two eyes. The trends are rather clear-cut but the particular points determining the functions are rather erratic, particularly as between the in- and out-of-phase curves. One uncontrolled source of instability lies in the fact that the in- and out-of-phase curves had to be determined at different sessions, thus confounding comparisons of effects due to phase of flash with day to day or time of day variance, with variations due to changes in subjective criterion of flicker adopted by *S*, and perhaps with effects of slight differences in alignment of the pupils of the eyes with respect to the oculars from one time to another. Variability within the five settings determining any one point tend to be small, *e.g.* standard deviations in general are less than one flash per second except at the highest intensities. On the other hand, variations in the means for any one point from one session to another are in general much greater, especially at the higher levels of luminance. Thus the differences pictured in Figs. 6 and 7, especially those representing different sessions are very likely to be unstable in detail.

The results certainly indicate interaction at a sensory level between the inputs from homonymous hemiretinas. Sherrington is undoubtedly correct, however, in his conclusion that excitations from corresponding points do not interact in a manner analogous to the integrative mechanisms at the entrance to motor final common paths.¹⁰ The small magnitudes of the summative and subtractive effects reflect at most only partial interactions.

Binocular interactions which occur under conditions of interocular differences of stimulation in which fusion is still possible (thus ruling out rivalry) all tend to involve a compromise sensory effect in which the perceived value of the stimulus is intermediate to the unocular components observed alone. Thus Ogle says,

Whenever disparate images fall within Panum's area and are fused and therefore are seen in one subjective visual direction . . . two possibilities exist: (1) both lose their identity in favor of an emergent compromise visual direction that lies between the two primary directions; or (2) the subjective visual direction of the fused image coincides with the primary direction of one of the eyes, while the other loses its identity entirely.¹¹

Apparently both possibilities occur: the former more often but the latter

¹⁰ Sherrington, *op. cit.*, 383.

¹¹ Ogle, *op. cit.*, 86-87.

can occur, especially in persons having marked ocular dominance. A binocular image appears intermediate in brightness when combined from unocular stimuli which differs in luminance.¹² Unocular stimuli of different color can combine binocularly to yield an intermediate hue.¹³ Similarly, a dim light in one eye lowers the CFF of a more intense stimulus on the corresponding retina of the opposite eye. On the other hand, straightforward summation can occur as when synchronous flashes in the two eyes yield higher critical flash rates than those of either eye alone. There is indeed interdependent functioning of the two eyes. The relative magnitude of all these effects tends, however, to be very unimpressive compared to the result of combining the stimuli on the same region of one retina.

The results of this study and of the others cited pose something of a paradox. On the assumption of some doctrine like 'isomorphism'¹⁴ or Müller's psychophysical axioms as operationally interpreted by Boring,¹⁵ the demonstrable behavioral facts of binocular fusion and interdependent function of processes from homonymous hemiretinas should reflect a 'two in, one out' integrative mechanism coupling the input from the two eyes. Either binocular fusion is epiphenomenal, *i.e.* not reflecting any underlying neurophysiological interactions (as implied in Sherrington's characterization of it as 'psychic'), or we have need to postulate (or better, to discover) a neurophysiological mechanism that achieves unification of converging processes without marked summative or inhibitive interaction. The evidence indicates there is some interaction of processes from corresponding areas of each retina as reflected in the final visual effect. One might suggest there is only partial confluence of the independent excitations from homonymous hemiretinas on cortical centers common to both. Perhaps only part of the total excitation from crossed and uncrossed visual pathways converges on common cortical cells, like the isodynamic cells of Cajal, and their activity is the basis of the relatively small amount of binocular interdependency. The rôle of such partial integration in the overall functioning of binocular vision is not clear, but it certainly appears to be one of the characteristics of the mechanism coupling excitation from corresponding regions of the two retinas into the cortical cyclopean retina.

¹² De Silva and Bartley, *op. cit.*, 540-541.

¹³ Selig Hecht, On binocular fusion of colors and its relation to theories of color vision, *Proc. Nat. Acad. Sci.*, 15, 1932, 369-389.

¹⁴ Wolfgang Köhler, *Gestalt Psychology*, 1947, 151 f.; cf Kurt Koffka, *Principles of Gestalt Psychology*, 1935, 56 and 62-66.

¹⁵ E. G. Boring, An operational restatement of G. E. Müller's psychophysical axioms, *Psychol. Rev.*, 48, 1941, 457-464.

SUMMARY

Sherrington is undoubtedly correct in his conclusion that fusion of two uniocular images in binocular perception is not based on an integrative mechanism analogous to the integrative mechanisms at the entrance to final common paths in the motor system. On the other hand, the evidence clearly points to some degree of interocular interdependency. Experiments were performed to determine the effect of interocular differences in intensity and phase of flash on *CFF* with binocular regard. The stimulus in the right eye was constant at an intermediate level of luminance for each experiment and varied over about 6 log units (millilamberts) in the left eye. *CFF* for the right eye alone at the intermediate luminance, for the left eye alone, and for both eyes at various levels of luminance of the left stimulus were determined with two Ss. These measures were made with both in- and out-of-phase flashes in the two eyes, with stimuli of 6° and 2.2° angular subtense, and at two retinal locations, central and 10° peripherally from the fovea. Results indicate a general damping or subtractive effect of a less intense stimulus in one eye on the *CFF* of a brighter stimulus on the corresponding area of the other retina—except with the smaller target at lower levels of luminance. As the uniocular stimuli approach equal luminance, in-phase, flashes yield summation, *i.e.* the *CFF* for both eyes is slightly higher than with out-of-phase flashes or with either eye alone. The summation of in-phase flashes is apparent even when the flash-rate is still above the flicker threshold of the less intensely illuminated eye, suggesting that binocular interaction occurs before flicker fusion in the chain of excitation in the visual pathways. Because of the relatively small magnitude of the interocular effects (compared with the effect of combining the stimuli in one retina) these results are interpreted as indicating only a partial convergence of excitation from corresponding points in the two retinas on cells in the visual cortex common to both the crossed and uncrossed pathways, *i.e.* the isodynamic cells of Cajal. Such partial integration seems characteristic of the sensory coupling of homonymous hemiretinas, but its rôle in binocular perception is unknown.

TEMPORAL CHANGES IN THE MEMORY OF VISUALLY PERCEIVED FORM

By JAMES C. CRUMBAUGH, University of Texas

The problem of temporally progressive changes in the memory of perceived form, as conceived by Gestalt theory, dates from an experiment by Friedrich Wulf in 1922.¹ There were several earlier experiments by Philippe,² Hempstead,³ Kuhlmann,⁴ and Crosland,⁵ but these researchers did not set out to examine the validity of Gestalt concepts as have the investigators since Wulf.

Wulf, approaching the problem from the orientation of the Gestalt laws of organization, called into question G. E. Müller's theory of "reactive forgetting." According to Müller's theory, which would, in general, probably find wide acceptance today, the memory-trace weakens or disintegrates with time, losing the distinctive characteristics of the original figure, except that any feature which is emphasized in perception will persist longer than others and will be accentuated in the reproduction.⁶ Wulf considered the possibility that memory might be governed by the same laws as those found by Wertheimer to hold for perception, the chief law being that of *Prägnanz*. The concept that every Gestalt becomes as 'good' as possible leads to the prediction that a figure reproduced from memory will be 'better' or 'more pregnant' than the figure as originally seen, and that the changes will be progressive with time.

Wulf presented a series of drawn figures to his *Os* and tested their recall of these figures by the *method of successive reproductions*. The chief characteristic of this method is that *O* makes repeated attempts to recall the original stimulus-figure at various intervals after its presentation, and changes in the recalled material are compared at each successive stage with the original figure. Wulf found the changes which occurred to be progressive in any one individual, sometimes in the direction of *leveling* and sometimes in the direction of *sharpening*. Leveling—the

* Accepted for publication September 6, 1953. This paper is adapted from a dissertation submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Psychology at the University of Texas, June, 1953. This study was directed by Professor M. E. Bitterman.

¹ Friedrich Wulf, *Beiträge zur Psychologie der Gestalt; Über die Veränderung von Vorstellungen (Gedächtnis und Gestalt)*, *Psychol. Forsch.*, 1, 1922, 333-373.

² J. Philippe, *Sur les transformations de nos images mentales*, *Rev. Phil.*, 43, 1897, 481-493.

³ L. Hempstead, The perception of visual form, this JOURNAL, 12, 1901, 185-192.

⁴ Frederick Kuhlmann, On the analysis of memory consciousness: A study in mental imagery and memory of meaningless visual forms, *Psychol. Rev.*, 13, 1906, 316-348.

⁵ H. R. Crosland, A qualitative analysis of the process of forgetting, *Psychol. Monogr.*, 29, 1921, No. 130, 1-157.

⁶ R. S. Woodworth, *Experimental Psychology*, 1938, 80.

disappearance of some sharply accented feature of the object, such as an asymmetrical jag in an arrow-figure—may result either from normalization or from autonomous change. Sharpening represents the opposite process of emphasizing an already accented feature.

Gordon Allport⁷ and F. T. Perkins⁸ produced results similar to Wulf's, as did Bartlett,⁹ whose study provides some of the most striking examples of these progressive changes. All of the investigators mentioned used the method of *successive* reproductions except Bartlett, who used the method of *serial* reproductions. In the latter method, the first *O* reproduces the stimulus-object from memory, then the second *O* uses this reproduction as a stimulus-object and reproduces it from memory after a time-interval, then the third *O* uses the second's reproduction, and so forth. Bartlett started with a stylized figure of an owl and found progressive normalization or assimilation toward a conventional figure, in this case a cat.

Another experiment which was interpreted as favorable to Gestalt hypotheses is that of Irwin and Seidenfeld,¹⁰ who used the method of *successive comparisons* and tested recognition at the various time-intervals by the method of *identical stimuli*—developed for the study of time-errors.

A serious error is inherent in the method of *successive* comparisons or reproductions (involving either recognition or recall of the stimulus-object). The successive methods offer an opportunity for fresh learning at each testing of memory—the trace of the original stimulus is influenced by *O*'s reproduction or comparison. This procedure obscures any true curve of forgetting that may exist; in fact, it favors consolidation rather than disintegration. As for the method of *serial* reproductions used by Bartlett, it does not measure the progress of the memory-trace with time; it reflects only whatever tendencies toward stereotypy may be common to a group of *O*s.

Gibson,¹¹ using the method of *immediate* reproduction, exposed a number of visual figures in quick succession and classified the changes which occurred in reproduction. He found object-assimilation (which Wulf called "normalizing") to be strongly influenced by verbal factors, and he was critical of the Gestalt interpretation of these changes. Similar results were obtained by Carmichael, Hogan, and Walter,¹² who presented ambiguous figures to one group of *O*s with the suggestion that they resembled certain objects and to another group with the suggestion that they resembled other objects.

Warner Brown was perhaps the first to study the progress of the memory-trace by the method of *single reproductions*,¹³ thus avoiding the error of *successive reproductions*. In the method of *single* reproductions the stimulus is presented only once to a given *O* and his memory for it is measured at only one time-interval. A differ-

⁷ G. W. Allport, Change and decay in the visual memory image, *Brit. J. Psychol.*, 21, 134-148.

⁸ F. T. Perkins, Symmetry in visual recall, this JOURNAL, 44, 1932, 473-490.

⁹ F. C. Bartlett, *Remembering*, 1932, 1-317.

¹⁰ F. W. Irwin and M. A. Seidenfeld, The application of the method of comparison to the problem of memory change, *J. Exper. Psychol.*, 21, 1937, 363-381.

¹¹ J. J. Gibson, The reproduction of visually perceived forms, *J. Exper. Psychol.*, 12, 1929, 1-39.

¹² Leonard Carmichael, H. P. Hogan, and A. A. Walter, An experimental study of the effect of language on the reproduction of visually perceived forms, *J. Exper. Psychol.*, 15, 1932, 73-86.

¹³ Warner Brown, Growth of 'memory images,' this JOURNAL, 47, 90-102.

ent group of *O*s is used for each time-interval, and there is thus no opportunity for new learning to contaminate the memory-trace. Brown found his results difficult to classify; some of his data were consistent with Gestalt theory and some did not seem to be. Goldmeier,¹⁴ using the method of *single reproductions*, measured the divergence of *O*'s drawings made after a given time-interval from the same *O*'s copy of the original stimulus-object made at the time it was presented. Goldmeier's results favored the Gestalt hypotheses, although Hebb and Foord have criticized both his procedure and the way in which he interpreted his data.¹⁵

Hanawalt,¹⁶ using Wulf's figures as stimulus-objects, compared the method of successive reproductions with the method of single reproductions. His over-all results were interpreted as unfavorable to Gestalt hypotheses. Hebb and Foord did work much like that of Hanawalt, avoiding successive comparisons,¹⁷ but where Hanawalt tested memory by the method of single reproduction (recall), Hebb and Foord used the method of single comparisons by multiple choice (recognition). They concluded that the results contradicted Gestalt theory. The most recent study in the field is another by Hanawalt¹⁸—a repetition of the experiment of Irwin and Seidenfeld using the method of *single* comparisons of identical stimuli instead of the method of *successive* comparisons of identical stimuli. On the whole, the experiment produced "little positive evidence concerning the nature of the trace and its relationship to the negative time-error."

From this summary of previous research on the problem of temporally progressive memory-change, two conclusions may be drawn: (1) There is great confusion and inconsistency of results in the literature; and (2) this confusion grows out of a failure to devise a clear-cut and universally acceptable methodology for attack upon the problem.

The principal issue may be stated briefly as follows: Does the memory-trace decay haphazardly with time, or are there lawful processes at work which influence its progress? Two alternatives to random disintegration can be considered—*autonomous* change and *assimilative* change. Thus the question becomes a two-fold one: (1) Do autonomous changes influence the progress of the trace; and (2) do assimilative changes influence its progress? The latter question has been satisfactorily handled in the literature; the work of Kuhlmann,¹⁹ Crosland,²⁰ Gibson,²¹ Carmichael, Hogan,

¹⁴ Erich Goldmeier, Progressive changes in memory traces, this JOURNAL, 54, 1941, 490-503.

¹⁵ D. O. Hebb and E. N. Foord, Errors of visual recognition and the nature of the trace, *J. Exper. Psychol.*, 35, 1945, 335-348.

¹⁶ N. G. Hanawalt, Memory trace for figures in recall and recognition, *Arch. Psychol.*, 31, 1937, No. 216, 1-89.

¹⁷ Hebb and Foord, *op. cit.*, 335-348.

¹⁸ Hanawalt, The method of comparison applied to the problem of memory change, *J. Exper. Psychol.*, 43, 1952, 37-42.

¹⁹ Kuhlmann, *op. cit.*, 316-348.

²⁰ Crosland, *op. cit.*, 1-157.

²¹ Gibson, *op. cit.*, 1-39.

and Walter,²² and others gives a clearly affirmative answer, although there is no good evidence of *progressive* change. It is the former question to which the present experiments are directed.

METHOD

In the present investigation, the method of identical stimuli was employed. As in the traditional time-error experiments, *O* made repeated comparisons of standard (*S*) and variable (*V*) which were introduced on opposite sides of a fixation-point.

Observers. A total of 285 *O*s were tested individually in the five experiments performed. All *O*s were drawn as volunteers from the psychology library at the University of Texas during the spring and summer of 1953. There were about

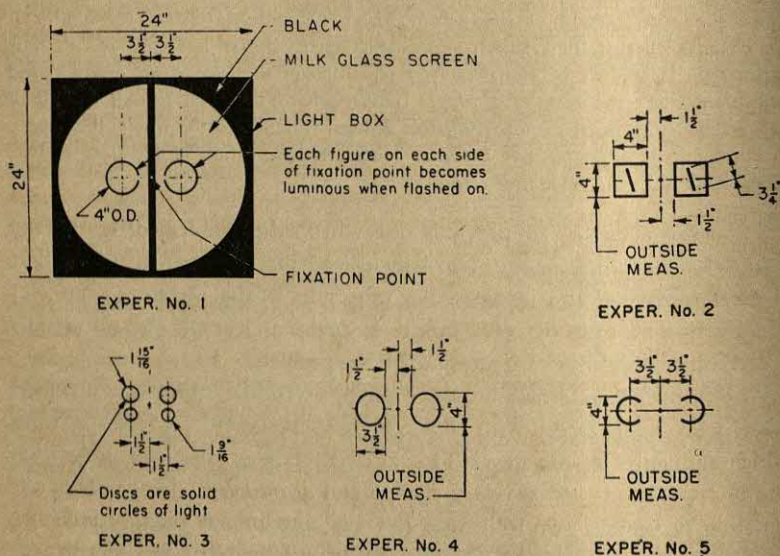


FIG. 1. DIAGRAMS OF THE STIMULUS-FIGURES AS THEY APPEARED IN THE LIGHT-BOX

twice as many men as women *O*s, chiefly because that was the distribution among the students in the library.

Stimuli. The stimulus-objects consisted of five geometric figures, each involving some type of asymmetry (see Fig. 1): (1) a broken circle (gap subtending 15° of arc); (2) a tilted line enclosed in a square (line rotated 15° counterclockwise from the vertical); (3) two disks with a size-ratio of $25/31$, the larger directly above the smaller; (4) an ellipse (ratio of diameters = $7/8$); and (5) a broken circle (90° gap). All figures, which were cut from black paper and mounted on the inside of

²² Carmichael, Hogan, and Walter, *op. cit.*, 73-86.

a milk-glass screen in a double light-box, appeared luminous on a dark background. Five experiments were performed, each utilizing one of the figures.

Apparatus. The apparatus consisted chiefly of a double light-box which was approximately a 24-in. cube, resting on a table 30 in. high. It was partitioned into two light-compartments, and the front had a narrow wooden strip at the dividing line between the two halves on which was mounted the fixation-light, a GE NE-2 neon glow-tube masked to a 1/8-in. aperture. The exterior of the light box was black and the interior was white. Each compartment was illuminated by a 25-w., 110-v. clear-glass lamp. In the front of the box was a circular milk-glass screen, half of which was illuminated from one compartment and half from the other.

E sat behind the light-box. A small 7-w., 110-v. shielded pilot-lamp furnished illumination for manipulating the control-switches and recording the results. To *E*'s left were a Hunter interval-timer, Model 111A, which timed the intervals between *S* and *V*, and two holding-relays which controlled the exposure-time (1/2 sec.) for both *S* and *V*. A 10-w., 110-v. lamp, shielded from *O* by a gooseneck reflector, dimly illuminated the experimental room, thus serving to structure the perceptual situation. The *O* sat in a chair at a distance of 7 ft. 7 in. from the fixation-point.

Procedure. *O* was seated and given standard instructions for the particular geometrical figure with which he was being tested. (Each *O* was tested with only one figure.) *O* was told that this was an experiment in the perception of visual form, and that he was not in competition with other *O*s in any way. The task was to compare visual forms which would be flashed on the screen in pairs, one after the other. Sometimes the first member of the pairs would appear on the right half and sometimes on the left half of the screen. Each figure would be flashed on for only 1/2 sec. The asymmetry (tilt, gap, or so forth) of the second figure would always be slightly different from that of the first figure. If the asymmetry of the second figure was greater than that of the first, *O* was to respond 'Greater'; if the asymmetry was less, *O* was to respond 'Less.' If *O* could see no difference in asymmetry he was to respond 'Equal' (in which case the same pair was repeated until he could give a definite judgment). From the signal 'Ready' until both figures of a pair had been exposed, *O* was to look steadily at the fixation-point. A sample trial was given (with an interval between *S* and *V* not used in the experiment). Neither on this trial nor at any time was *O* given any indication of the 'correctness' of his response. Furthermore, *O* was never given any explanation of the true nature of the experiment, even after he had been tested, a procedure adopted to prevent the information from reaching the other potential *O*s in the library.

The time-intervals between *S* and *V* which were employed in the different experiments were 0, 0.3, 2.0, 5.0 and 12.0 sec. At 0 sec. the onset of *V* coincided with the offset of *S*. In the first three parts of Experiment I (Ia, Ib, Ic) only certain of these intervals were used, and in the special case of Ib the two stimuli were presented *simultaneously* as well as at certain positive intervals. In Experiment Id, and in all subsequent experiments (II, III, IV, V) all five of these time-intervals were utilized. All stimulus-figures (both *S* and *V*) were exposed for 1/2 sec.

The various time-intervals were scheduled according to a 'planned haphazard order' in which each interval would appear equally often in each position in the series of five intervals. The order employed for the first half of the trials was

replicated in reverse for the last half. In all experiments except in the first three parts of Experiment I, each interval was repeated eight times, making a total of 40 judgments for each *O*. In subexperiments Ia, Ib, and Ic, each interval was repeated 10 times. Ia utilized four intervals (0.3, 2.0, 5.0, 12.0 sec.) making a total of 40 judgments per *O*; Ib utilized three intervals (simultaneous, 0, 0.3), making a total of 30 judgments per *O*; and Ic utilized three intervals (0, 0.3, 5.0) making a total of 30 judgments per *O*. These three parts of Experiment I required a somewhat different order for presenting the various time-intervals, but the schedule was in general quite similar to that employed in the remainder of the investigation. In all experiments, and for each time-interval employed in a given experiment, *S* was presented an equal number of times on the right and on the left side of the screen (with *V* always presented on the side opposite *S*).

RESULTS

Experiment I. A broken circle with a gap subtending 15° of arc (Fig. 1) was the first figure to be studied because it is one of the simplest possible asymmetrical figures. The orthodox Gestalt laws of organization predict a clear-cut change in the direction of increased symmetry (closure of gap) with time. It was necessary to do considerable exploratory work (subexperiments Ia, Ib, Ic) to determine the time-intervals between *S* and *V* most suitable for demonstrating this change, and all subsequent work was done with these intervals. The results of all experimentation with the circle containing a gap of 15° are reported graphically in Fig. 2 in terms of the percentage of *judgments of increasing symmetry of the trace* with time. An increase in the percentage of such judgments does not necessarily reflect an increase in the *degree* of symmetry with time, although it seems reasonable to assume such a relationship (as is commonly done).

Inspection of Fig. 2 shows a general tendency for the gap first to increase and then to decrease. The process is progressive in time, although its ultimate fate cannot be determined from the present data since the longest interval (12.0 sec.) still shows evidence of change in progress. It is impossible to say whether the curve would reach the degree of symmetry existing in the original figure (*S*) at approximately 13 sec. (as extrapolation would indicate), or whether it would go on to a degree of symmetry greater than that of the original figure. Orthodox Gestalt psychology might predict the latter result.

Wilcoxon's test for paired replicates was selected for the statistical evaluation of the results.²³ This non-parametric technique provides a more conservative estimate of significance than does the *t*-test. The pooled results

²³ Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures*, American Cyanamid Co., Stamford, Conn., 1949, 1-16.

for all parts of Experiment I could not be analyzed since different time-intervals were studied in the different parts. It is obvious, however, that the results for the various subexperiments are quite consistent. The decline between 0 and 0.3 in Experiment Ic is significant beyond the 5-% level, and the corresponding decline in Id is significant at the 1-% level. The increase between 5.0 and 12.0 sec. is significant at the 5-% level in Experiment Ia and at the 1-% level in Id.

In an effort to determine whether changes of the kind which appeared in Experiment I are general or peculiar to a given figure, a second experiment utilizing a different figure was conducted. A tilted line was selected,

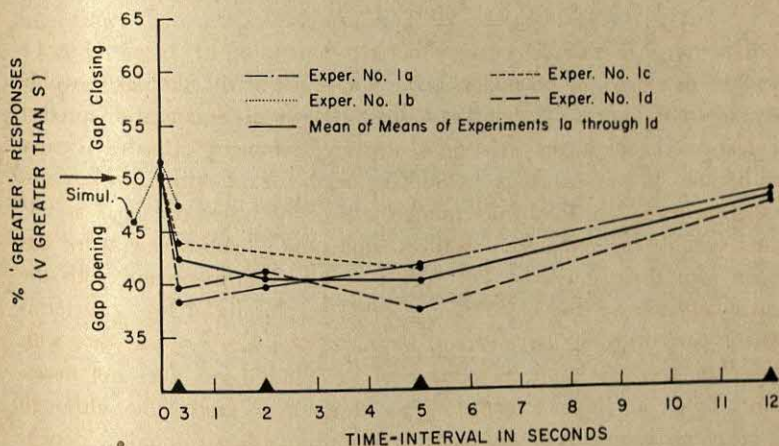


FIG. 2. RESULTS OF EXPERIMENT I

which, like the broken circle, is a simple figure for which orthodox Gestalt theory would predict increasing symmetry.

Experiment II. The data of this experiment, in which a total of 35 Os were tested, are plotted in Fig. 3. The form of the curve for the tilted line closely parallels that for the circle of Experiment I. Wilcoxon's test for paired replicates, when applied to the differences between the high and low points of the curve, shows significance at the 1-% level. The next greatest difference is between the intervals of 0 and 2.0 sec. which is significant at approximately the 3-% level.

The question arises as to whether the results for both figures might be artifacts of a judgmental or 'guessing' pattern common to the majority of

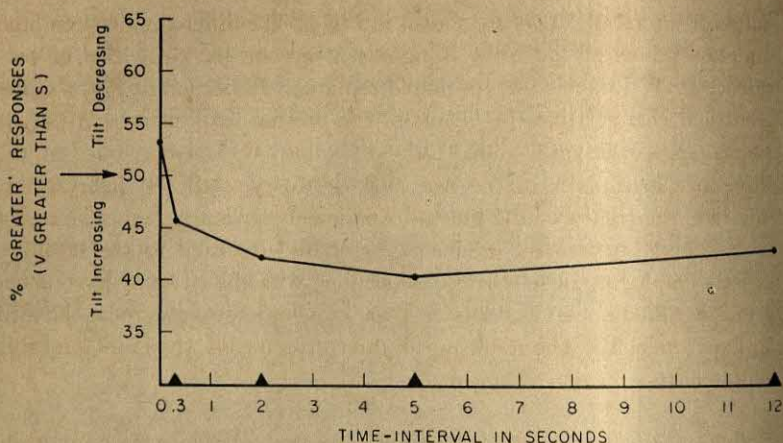


FIG. 3. RESULTS OF EXPERIMENT II

Os. To answer this question, Experiment III was designed. It utilized a quite unorthodox figure (consisting of two disks of different size) for which quite another pattern of change in time might be anticipated. If the same pattern of results were obtained, the previous results might be artifacts of the testing situation rather than true changes in the memory-trace.

Experiment III. The difference in size of two disks with a ratio of 25/31 in diameters was judged by *O* as greater or less in *V* than in *S*. The re-

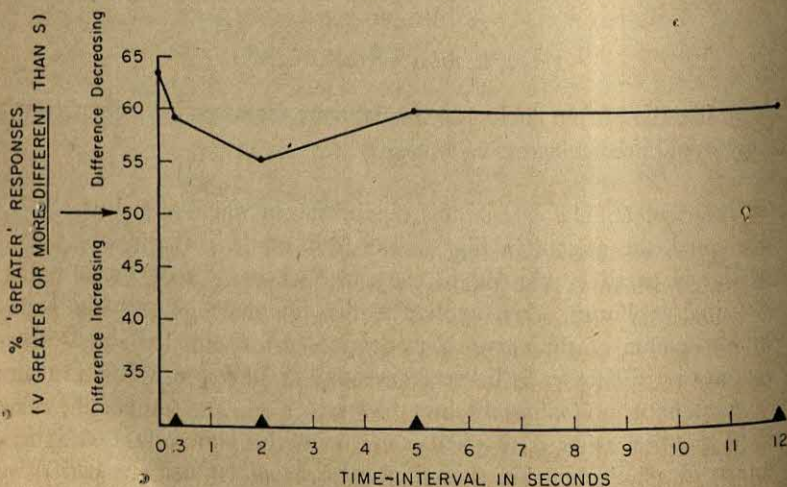


FIG. 4. RESULTS OF EXPERIMENT III

sults for a total of 50 *O*s are shown in Fig. 4. The difference between high and low points of the curve is significant beyond the 5-% level of confidence by Wilcoxon's test for paired replicates. Inspection of Fig. 3 shows a marked change from the results of the previous experiments. All of its points are consistently above the 50-% line—the memory-trace of the difference in the size of the two disks decreases at all *S-V* intervals. At any rate, the results of the present experiments obviously cannot be traced to a general pattern of judgment essentially unrelated to the stimulus-objects employed, since the results obtained with the circle of Experiment I and the tilted line of Experiment II are quite different from those obtained in Experiment III. The results of all three experiments, therefore, probably represent true memory-changes.

Experiment IV. An ellipse was chosen for this experiment because, while quite different from previously used figures, it involves a simple asym-

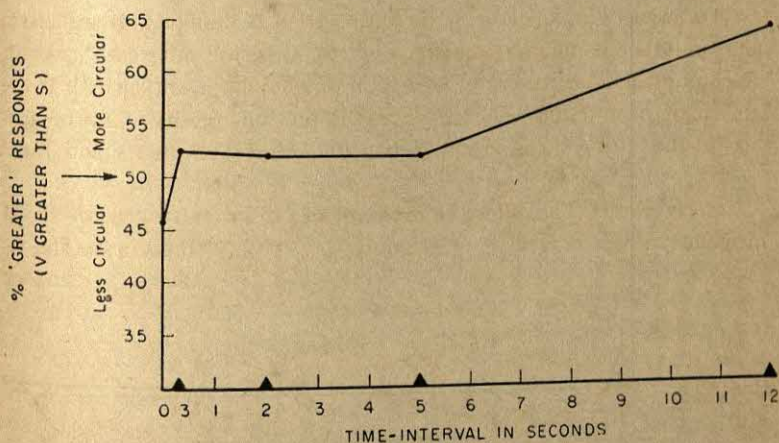


FIG. 5. RESULTS OF EXPERIMENT IV

metry that would be expected from the Gestalt laws of organization to decrease with time. The results, in which a total of 30 *O*s were tested, are shown in Fig. 5. The difference between the high and low points of the curve is significant beyond the 1-% level of confidence by Wilcoxon's test for paired replicates. The next greatest difference, between 5.0 and 12.0 sec., also is significant beyond the 1-% level. The form of the curve is qualitatively different from that of any of the previously used figures, which confirms the suggestion from Experiment III that the pattern of memory changes to be expected in the trace of a visual form is a function

of the form involved. The trace of the ellipse shows no significant initial change in symmetry of figure followed by an increase in symmetry beyond that of *S*.

It seems, then, that a qualitative change in the originally perceived figure may yield a qualitative change in the form of the curve representing change in the memory-trace with time, although some qualitatively different figures have common patterns of change. The question arises as to whether a *quantitative* change in the originally perceived figure would yield *quantitatively* different patterns of change (similar in pattern but displaced in degree), or whether the changes would be qualitatively different. To answer this question Experiment V, utilizing a broken circle with a gap-size different from that of Experiment I, was performed.

Experiment V. The circle with gap subtending 90° of arc was studied with 40 Os. The results are shown in Fig. 6. The difference between the

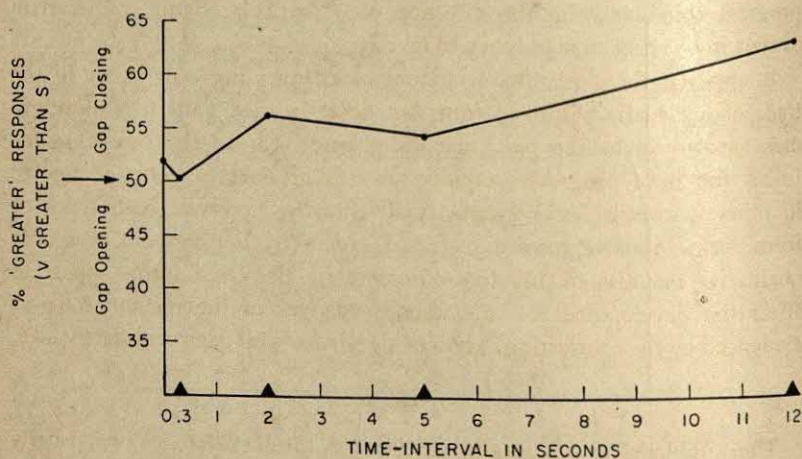


FIG. 6. RESULTS OF EXPERIMENT V

high and low points of the curve is significant beyond the 1-% level of confidence by Wilcoxon's test for paired replicates. The next greatest difference, between 5.0 and 12.0 sec., is significant at approximately the 3-% level. The curve of Fig. 6 undoubtedly did not arise by chance. It represents a general increase in the symmetry of the trace with time, and would seem to represent a different pattern than in the case of the 15° gap. In a certain sense, then, a quantitative change in the original figure may be said to have produced a qualitative change in the pattern

of progress of the trace—an increase in the size of the gap has eliminated the initial increase in asymmetry.

In view of the results of all experiments it seems clear that symmetry of figure (autonomous change) rather than assimilation to some standardized or normative figure is the chief factor involved in the obtained progressive changes. It is difficult to see, for example, what normative figure could be involved in the increasing tilt of a line in a square. Similarly, a change (in either direction) in the relative size of two disks hardly appears to improve their normative value, and it is illogical to assume that a broken circle which becomes *more* broken (increase in the 15°-gap at the shorter intervals) is being assimilated to a standardized concept of 'circle.' It is still more difficult to assume that most people have a normative concept of '*broken* circle.' For these reasons the present results may be interpreted as favoring the concept of autonomous change. If the changes were assimilative, we should certainly expect them to progress consistently in the direction of a particular figure, rather than to start in one direction and then to reverse.

It appears, therefore, that autonomous changes not only occur in the trace of a visually perceived form but differ in their pattern of development according to the particular form perceived. In the case of some forms the trace progresses rapidly towards greater symmetry of figure; in others, towards lesser symmetry. Eventually, however, both types of form seem to move toward increased symmetry. Furthermore, not only qualitative but also quantitative differences in perceived forms appear to affect the development of autonomous changes in the memory-trace as evidenced by the contrasting curves of the circles with gaps of different size.

SUMMARY

The problem of temporally progressive memory-changes in visually perceived forms was studied by the method of identical stimuli, in which identical figures (standard and variable) both were repeated in each comparison made by *O* as in a typical psychophysical experiment on the time-error. Luminous stimuli on a milk-glass screen were used, standard and variable being presented on opposite sides of a fixation-point. Each of the five figures employed contained a degree of asymmetry which *O* judged to be greater or less in the variable as compared with the standard.

The results provide clear evidence of temporally progressive changes in the traces of visually perceived forms. The changes observed, which suggest autonomous rather than assimilative effects, differ in their pattern

of development according to the particular form perceived. The traces of some forms progress rapidly toward greater symmetry, while the traces of others show an initial tendency toward decreased symmetry which soon is reversed. The direction and pattern of change is influenced by quantitative as well as by qualitative variations in the perceived figure. While autonomous changes as yet undiscovered may occur over longer intervals, the present experiment reveals important effects within the first few seconds of the life of the trace.

THE EFFECT OF SIZE OF TARGET ON ACCURACY OF AIM

By R. H. DAY, University of Bristol, England

Previous studies of aiming have yielded results which demonstrate that modifying the size of a target, either by changing the actual physical dimensions or by altering the distance of the target from *S*, results in variations in accuracy of aim.

Experiments by Mace,¹ Lewin and Voigt,² and Moffett³ have shown that there is a functional relationship between target-size and efficiency of performance where performance is judged in terms of the accuracy of aim. Lewin and Voigt were able to show that accuracy increased, within limits, with increase of the distance of the target from the *S*, and that accuracy depended to some extent upon the structure of the target or 'goal.'

Using Muscio's aiming-test and a dart-board test, Mace has shown that a 'good,' 'fair,' or 'poor' shot tended to be defined by *S* in a way which was relative to the form of target used. Generally, the greater the size of target the smaller the *relative* subjectively assessed error irrespective of the *absolute* distance from the bull's eye. In his interpretation of these facts Mace has introduced the notion of an "implicit standard," a subjective estimate of performance based upon the objective design and size of the target. The implicit standard Mace has defined as, "a level of attainment such that performance below this level is followed by stabilization (provided stabilisation is not accounted for by any 'physiological limit')." Generally, it seems that a modification in the design of the target entails a corresponding modification in the efficiency of performance through the operation of an implicit standard set up by the particular target-design. In concluding his report Mace posed the question regarding whether there is a certain limiting value beyond which the functional dependence of accuracy of aim on target-size fails to apply, or is reversed.

An attempt has been made in this investigation to answer the question posed by Mace by varying the size of the target over a moderately wide range while holding constant the essential design and structure of the target.

* Accepted for publication October 6, 1953. This research is part of a project carried out under contract A100624/51/C.10 between the University of Bristol and the Flying Personnel Research Committee of England. Grateful acknowledgement is made to Professor G. C. Drew and Mr. A. E. Earle for their helpful suggestions.

¹ C. A. Mace, The influence of indirect incentives upon the accuracy of skilled movements, *Brit. J. Psychol.*, 22, 1931, 101-114.

² Kurt Lewin and G. Voigt, Untersuchungen zur Handlungs- und Affekt-psychologie: XI. Über die richtungspräzision linear fernhandlung, *Psychol. Forsch.*, 16, 1932, 70-113.

³ D. C. Moffett, A study of the accuracy of direction in motor skills as determined by the relative size of the angle of error. *Res. Quart. Amer. Ass. Hlth. Phys. Educ.*, 13, 1942, 466-480.

METHOD AND PROCEDURE

Subjects. Twenty-five Ss (university students) were used in the experiment. All were right-handed.

Apparatus. A plywood panel 38 cm. x 43 cm. was fixed to the wall by means of two bolts fitted with wing nuts and placed at a height of 149 cm. from the floor. This panel had two vertical slots through which the bolts projected thus making it possible to raise or lower the panel to suit S's height. Fixed to the panel was a sheet of cork 1 cm. thick and measuring 27 cm. x 28.5 cm. with clips above and below it, for fastening the target-sheets firmly over the cork surface. The plywood panel was painted dull black and the cork was unpainted.

The target-sheets were of white foolscap paper 33 cm. x 20.2 cm.; each sheet having 12 targets arranged in four rows of three. The distance between target-centers in the horizontal and vertical directions was 7 cm. for all sizes of target.

There were five sizes of target, the diameters of which were 1, 2, 3, 5, and 6 cm. Each target consisted of a bull's-eye 1 mm. in diameter and three concentric rings so arranged that the distance of the inner ring from the bull's eye and the distance between the three concentric rings was the same. The five targets are shown in Fig. 1. For convenience these targets in future will be referred to as Target 1, Target 2, etc. beginning with the smallest size. The targets employed in the preliminary matching trials consisted of crosses with the intersecting arms arranged diagonally relative to the sheet, each arm being 1.5 cm. long. The arrangement of these crosses on the sheet was the same as the targets used in the experiment proper.

The stylus with which Ss speared at the targets was 17 cm. long tapering at each end and 0.8 cm. in diameter. At the striking end was a steel tip 1 cm. long and tapering to a sharp point.

The rate at which S speared at the targets was controlled by the beat of a metronome which was set at 184 beats per minute throughout the experiment. In all trials the metronome was placed behind S.

Experimental design. The design of the experiment closely followed that suggested by Edwards in which Ss were matched across rows on the basis of performance on a similar task.⁴ Every S served twice, the sessions separated by a period of approximately one week.

In the first session S was required to aim at the centers of the crosses, 2 sets of 12 being used, in order to assess initial ability in this kind of task. This session also served to familiarize S with the general nature of the task and provide practice in spearing at the targets in strict time with the metronome. On the basis of performance in this initial task the Ss were ranked for ability and divided into five groups. Throughout the experiment the error-score was the distance from the center of the target to the point at which the stylus struck the paper. It was found possible to measure this distance to the nearest 0.5 mm. In dividing the Ss into groups the first group included the five Ss with the lowest mean error-score, the second group of five with the second lowest mean error-score and so on on to the last group, which included the five Ss with the highest mean error-score. Using a table of random numbers one member of each of these ability groups was assigned to one of the five targets so that these groups were roughly matched for initial ability.

⁴ A. S. Edwards, *Experimental Design in Psychological Research*, 1950, 264-283.

Procedure: Session 1. Before beginning the task the height of the panel was so adjusted that its horizontal mid-line was at the same height as S's hand held outstretched at shoulder-level. S stood squarely before the panel, and at arm's length from it. In this session the general nature of the task was explained and S was

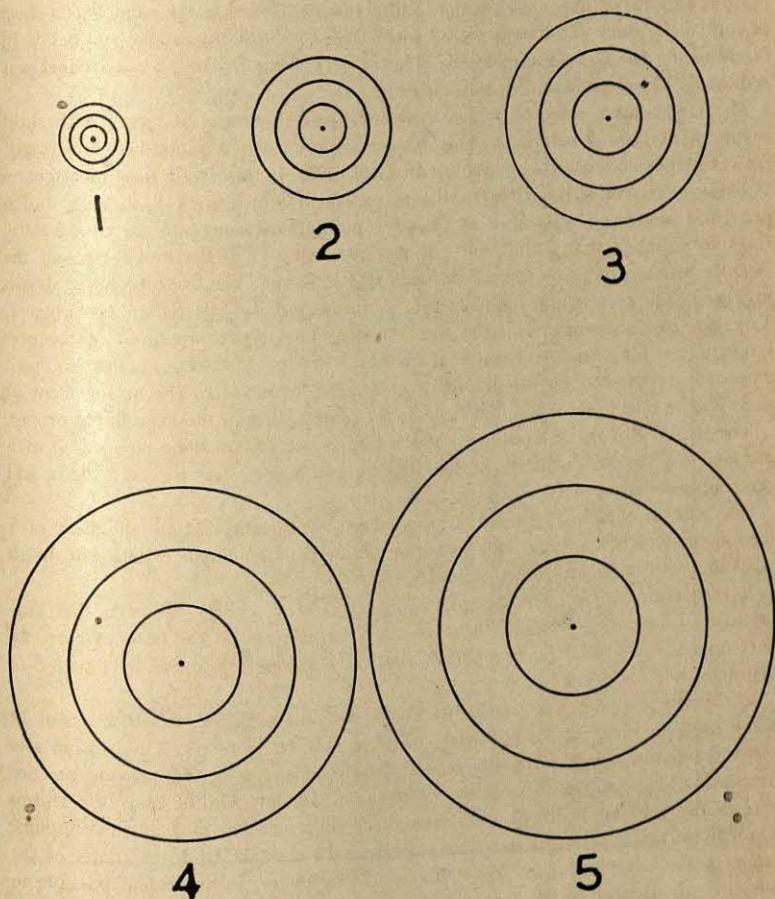


FIG. 1. TARGETS USED IN THE STUDY

given practice in spearing at a blank sheet of paper in time with the beat of the metronome. In making these movements S speared at the target on one beat and withdrew his hand to the shoulder at the next thus completing the movement in 0.65 sec., and taking 7.8 sec. per set of 12 shots. After 3 min. of practice on a blank sheet of paper S was given 2 sets of 12 shots at the crosses. A standard

pattern of movement was followed on each target-sheet, the first shot being at the top left target, then along the top row from left to right followed by the second row beginning at the right, and finally completing the 12 shots on the bottom left target. It was made clear to *S* that he should keep in strict time with the beat of the metronome and that he should strike each cross at the point of inter-section of the arms.

Session 2. Before commencing the task in the second session the panel was again adjusted for height and *S* was placed in the correct position and asked to maintain that position throughout the session. During this session *S* was required to spear at five sets of targets with a rest-pause of 2 min. between the third and fourth sets. Before beginning the task the instructions were repeated and timing and accuracy again emphasized. *S* was told to commence spearing at the targets any time after the metronome was set in motion. After each set of 12 shots was completed a new sheet of targets was clipped to the panel and the performance repeated. The same pattern of movement was followed as in the preliminary matching-trials.

RESULTS

The mean error-scores for each set of 12 shots for each *S* are given in

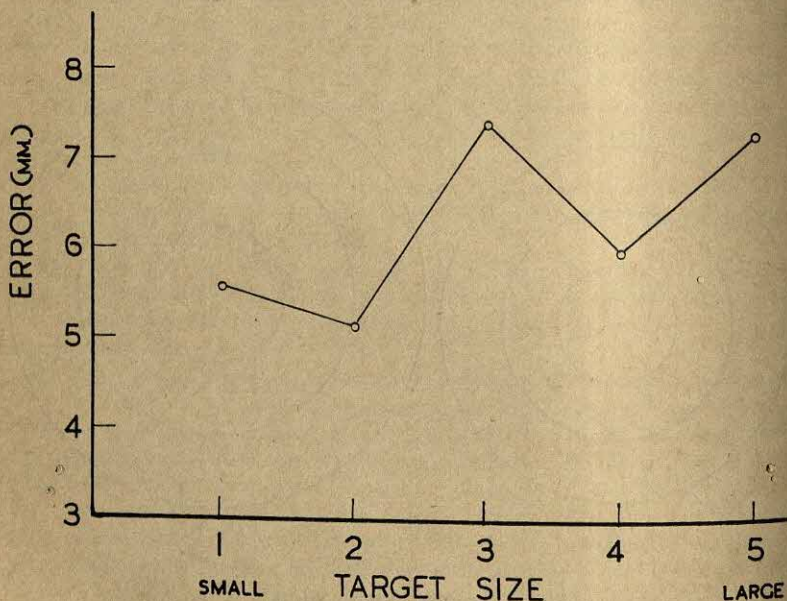


FIG. 2. VARIATION OF THE MEAN ERROR-SCORES AS A FUNCTION OF THE SIZE OF THE TARGET

Table I, and the means and standard deviations of these scores for the 5 target-sizes are shown in Table II. The general trend evidenced in Table II can be seen more clearly in Fig. 2, where it can be observed that there is

a noticeable variation in the mean error-scores as a function of target-diameter. There is a considerable increase in mean errors on Target 3 as compared with Targets 1 and 2 between which there is little difference. On Target 4 there is a tendency for the error to decrease followed by an increase on the largest target.

TABLE I
MEAN ERROR-SCORE FOR EACH S ON 5 SETS OF 12
SHOTS FOR EVERY TARGET-SIZE

Target-size	S	Sets of 12 shots				
		1st	2nd	3rd	4th	5th
1	1	3.58	5.25	5.45	5.08	3.50
	2	3.33	4.50	6.58	3.46	5.75
	3	8.66	4.91	5.45	7.54	6.83
	4	7.08	7.54	6.62	5.08	5.75
	5	5.91	8.17	4.62	4.62	4.54
	Mean	5.71	6.07	5.74	5.15	5.27
2	1	4.00	4.66	3.91	4.00	3.87
	2	4.30	4.12	5.12	3.57	3.62
	3	3.33	4.62	5.66	4.70	4.33
	4	6.62	5.12	6.16	5.00	3.54
	5	7.87	8.50	6.75	8.66	6.62
	Mean	5.22	5.40	5.52	5.24	4.39
3	1	10.58	7.41	7.95	7.75	7.08
	2	4.54	5.00	6.10	7.91	7.12
	3	8.72	7.95	7.70	6.62	6.12
	4	8.20	8.20	5.54	5.50	4.50
	5	9.96	9.08	11.44	8.58	9.09
	Mean	8.40	7.52	7.40	6.67	6.79
4	1	7.04	5.20	5.70	3.91	3.45
	2	4.08	5.20	3.12	4.66	3.37
	3	4.80	5.46	4.91	5.70	4.66
	4	7.70	8.33	7.83	6.16	5.54
	5	8.62	8.30	6.70	5.41	7.29
	Mean	6.45	6.49	5.65	5.17	5.26
5	1	6.37	5.54	7.16	7.30	5.41
	2	4.54	6.30	4.54	5.25	5.08
	3	10.37	9.37	7.75	4.50	9.41
	4	11.66	7.25	7.25	6.95	7.75
	5	8.33	7.46	5.96	9.79	5.89
	Mean	8.25	7.18	6.53	7.75	6.70
Mean for sets of 12 shots		6.80	6.53	6.23	6.00	5.68

An analysis of variance indicates that the variation in mean error-score as a function of size of target is significant beyond the 1-% level of confidence. A closer scrutiny of the significance of difference between means using the *t*-test has revealed that there is no significant difference between

error-scores on Targets 1 and 2 but the difference in performance on Targets 2 and 3 is significant at the 1-% level. The differences in mean error-scores for Targets 3 and 4, and for Targets 4 and 5 just fail to reach

TABLE II
MEAN ERROR-SCORE AND STANDARD DEVIATION FOR FIVE TARGET-SIZES

	Target-size				
	1	2	3	4	5
Mean error-score	5.58	5.15	7.41	6.00	7.28
Standard deviation	0.85	1.30	1.63	0.97	1.43

the 5-% level although there is a strong tendency for the error to decrease on Target 4 as compared with Targets 3 and 5.

To examine more closely the general trend of performance over the 5 sets of 12 shots for each target, the mean error-score for each set as shown in Table 1 has been plotted. These plots can be seen in Fig. 3. This graph

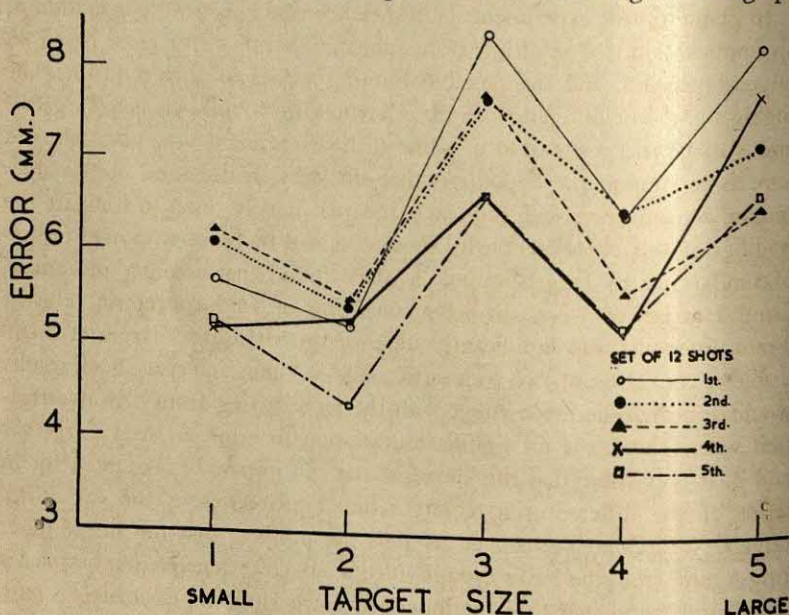


FIG. 3. MEAN ERROR-SCORES OF EACH OF THE 5 SETS OF 12 SHOTS AT EVERY SIZE OF TARGET

rather strikingly reveals that the trend evidenced in Fig. 2 has persisted and there is a marked consistency in the data for each set of 12 shots for the 5 target-sizes. Each plot shows a marked increase in error from Target 2 to Target 3, a less marked decrease from Target 3 to 4, followed by an

increase in error on Target 5. In no case is there a significant variation in error between the two smaller targets. In examining the differences between means with the *t*-test it has been found that in all sets of 12 shots the differences in mean error between Targets 2 and 3 were significant at the 1-% level. The difference between mean error-score for Targets 1 and 2 did not approach significance for any set of 12 shots. In the case of Targets 3 and 4 the difference between means was significant at the 5-% level for the first and third set of 12 shots. On the fourth set the difference was at the 5-% level for Targets 4 and 5.

In examining the successive sets of 12 shots without regard to target-size it can be observed that there was a small improvement in error-score from one set to the following set (Table I). This decrease, indicating a practice effect, did not satisfy the criteria for significance.

DISCUSSION

In planning this experiment interest centered mainly on the operation of an implicit standard resulting from subjective estimates of error based on target-dimension, and the possible limiting values of such a factor. The highly significant differences in mean error-score found between Targets 2 and 3 make the postulation of some such determinant very plausible and appears to support the hypothesis that efficiency in this kind of task does in part depend upon such implicit standards. It is in order to compare the findings of this experiment with those obtained by Mace who used targets of similar size to Targets 1 and 2. Mace found that accuracy in aiming, using a target of 1-cm. diameter consisting of four concentric rings at 1-mm. intervals, was significantly superior to that using a target of 2-cm. diameter consisting of two concentric rings at 5-mm. intervals. Such results would appear to be inconsistent with the facts arising from this investigation where there was no significant variation in error between Targets 1 and 2. It is possible that this inconsistency can in part be accounted for in terms of the differences in configurational properties of the two 2-cm. targets. With Mace's 2-cm. target it is not unlikely that the single inner ring 5 mm. from the bull's eye provided a less clear demarcation between a 'good' and 'poor' shot than the inner concentric rings approximately 3 mm. apart used in this experiment. Moreover, Mace's Ss performed this task, consisting of 150 shots, twice a day for 5 days, as compared with a short single session of 60 shots used in this experiment. A prolongation of the task could perhaps contribute towards differences in aiming accuracy not revealed in a short trial where the early stages of learning only came under

review. It can be reasonably supposed that such differences in the configurational properties of the target and the technique of experiment would account for the inconsistency noted above.

More important, however, is the tendency of the error to decrease on Target 4 as compared with Target 3, and to increase again on Target 5. The decrease is a reversal of the effect to be expected from Mace's hypothesis. Apparently the upper limit of the range over which the implicit standard operates has been reached with Target 3 and some other factor has operated to produce the decrease. This decrease, it will be remembered, was significant with Sets 1 and 3 and approached significance on the remaining sets. A possible clue to the cause of this decline in error is provided by investigations of check reading dials under a various conditions of dial-diameter. White has shown that the efficiency with which dials of 1.0, 1.75, and 2.75-in. diameter were check-read followed a trend similar to the trend in the efficiency of aiming at targets of 3, 5, and 6-cm. diameter.⁵ The three dial-diameters closely correspond to the target-diameters used in this experiment. White found that the elements determining the response-times and error-scores on the clock-reading task appeared to be the patterns of eye-movements which were slightly more efficient in terms of number and duration of fixations for the 1.75-in. dial than for the larger or smaller dial. Since Target 4 closely approximates in diameter the 1.75-in. dial some such ocular-motor determinant may well have been responsible for the slightly superior performance on Target 4. It is possible, then, that the trend in accuracy of aim for the three larger targets could be accounted for in terms of ocular-motor function rather than in terms of changes in implicit standards as a function of target-size. It could be supposed that phenomena such as these relating to the visual musculature and to the mechanics of eye-movements combined with the operation of an implicit standard to produce the differences in aiming accuracy found between Targets 2 and 3. On the basis of Mace's work, however, an explanation which emphasizes the part played by implicit standard would appear to be the more reasonable assumption since there is little evidence to support the first view.

In conclusion it is worthwhile to introduce into the discussion a useful notion of Helson which he has called "the hypothesis of par or tolerance."⁶

⁵ W. J. White, The effect of dial-diameter on ocular movements and accuracy of check reading groups of simulated engine instruments, *U.S.A.F. Air Material Command Technical Report 5826*, 1949.

⁶ Harry Helson, Design of equipment and optimal human operation, this JOURNAL, 62, 1949, 33-44.

As a result of his investigations of equipment-design, Helson has suggested that operators appear to have a standard or par of excellence representing the greatest possible error they will tolerate under certain conditions. This standard is a resultant of many factors including those of a motivational nature. It can be said that the variation in error between Targets 2 and 3 represents changes in the par of excellence brought about by a modification in the subjective assessment of error resulting from target-size. The variation in performance between the three larger targets, on the other hand could be due to variation in the par of excellence springing from more ocular-motor phenomena, the conditions of Target 4 favoring optimal performance.

SUMMARY

(1) An experiment has been carried out in which 25 Ss were required to spear at a series of targets in order to test the hypothesis that accuracy of aim is in part a function of target-size, and to find the limiting values, if any, of such a function.

(2) It has been shown that accuracy of aim considerably increased using a target of 3 cm. diameter as compared with targets of 1- and 2-cm. diameter between which there was little difference in accuracy of aim. On a target of 5-cm. diameter there was a nearly significant decrease in error as compared with Targets of 3 and 6 cm.

(3) The trend revealed by 5 sets of 12 shots for the 5 target-sizes showed a marked consistency with trend evidenced by the mean error for the 5 sets taken together.

(4) An hypothesis in terms of implicit incentives for the difference in error between Targets 2 and 3, and in terms of Targets 4 and 5 has been suggested.

(5) The concept of par of excellence suggested by Helson has been introduced as a useful and more unifying explanatory principle in the treatment of these data.

KINESTHETIC AFTER-EFFECTS FOLLOWING CEREBRAL LESIONS

By ROBERT JAFFE, Mount Sinai Hospital, New York City

The effect of prior stimulation in altering the apparent dimensions of objects held in the hand (kinesthetic figural after-effect) was demonstrated originally by Gibson.¹ His findings were extended by Köhler and Dinnerstein,² and in general confirmed in an exploratory study by the present author. In Köhler and Dinnerstein's study Ss rubbed the fingers of one hand along the sides of either wide or narrow strips of cardboard, for one or two minutes, and then used the same hand to judge the width of a standard strip of intermediate size. This procedure led to a systematic distortion in S's estimation of the width of the standard. After stimulation with the smaller strip, they judged the standard to be larger than its true size, and after stimulation with the larger strip, they judged the standard to be smaller than its true size. Köhler and Dinnerstein suggest that their results could ultimately be explained on the same basic premises of the satiation theory proposed by Köhler and Wallach to account for figural after-effects in vision.³

Köhler and Wallach have been specific in postulating that visual stimulation results in the formation of electrotonic fields in the cerebral cortex. Such fields are topologically related to the position of the stimulating object in the visual field. It is further assumed that these electrotonic fields induce changes in the 'conductivity' of the cortex surrounding the zone of primary cortical excitation. The induced changes in 'cortical conductivity' are postulated as the mechanism for the distortion in visual perception known as 'figural after-effects.' By contrast, Klein and Krech have ac-

* Accepted for prior publication September 13, 1954. This work was performed in the Psychophysiological Laboratory, Department of Psychiatry and Neurology, New York University College of Medicine. It was supported in part by a grant-in-aid from the Commonwealth Fund of New York and by Contract No. V 1001M-176 with the United States Veterans Administration. The author, Abrahamson Fellow, Department of Neurology, is indebted to Drs. Morris B. Bender, Hans-Lukas Teuber, and W. S. Battersby for their aid in this study.

¹ J. J. Gibson, Adaptation, after-effect, and contrast in the perception of curved lines, *J. Exper. Psychol.*, 16, 1933, 1-31.

² Wolfgang Köhler and Dorothy Dinnerstein, Figural after-effects in kinesthesia, *Misc. Psychologica Albert Michotte*, 1947, 196-220.

³ Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. Phil. Soc.*, 88, 1944, 269-357.

counted for such after-effects in terms of a less well-defined concept of generalized 'cortical conductivity'.⁴

The present study was undertaken to explore the adequacy of the satiation theory in dealing with kinesthetic after-effects. To this end, groups of normal and brain injured Ss were tested in a uniform manner. The underlying assumption was that the formation and distribution of 'electrical fields' in the cortex would be altered in patients with brain lesions, as compared with normal controls. If 'cortical conductivity' is indeed a factor in the formation of after-effects, then the amount and direction of figural after-effects in the brain-injured group should differ from those of the normal population.

METHOD AND PROCEDURE

Subjects. The experimental Ss were 20 veterans of World War II who had sustained penetrating brain injuries 2-7 yr. prior to testing. They were compared with 20 normal controls of the same sex and age-range. Precise localization of the total brain damage was impossible, but the lesions were judged to be unilateral and roughly localized on the basis of x-ray evidence of bone damage and retained foreign bodies, the notes of attending neurosurgeons, and the neurological signs and symptoms. The sensory function of the experimental Ss was evaluated in a comprehensive series of tests: touch perception on single and on double simultaneous stimulation, pin-prick and adaptation to pin-prick, vibration, passive movement, graphesthesia, stereognosis, and two-point discrimination. Ten of the Ss had impairment of one or more of the sensory functions tested. The other 10 had no sensory disturbance.

Materials. A strip of aluminum 2 in. wide was used as the standard. Two other strips were used for interpolated stimulation. One of these was 1 in. wide, and the other 4 in. The width judgments were made by comparison of the standard with a strip graduated in width; the width of this graduated strip varied, from 1 in., at one end to 3 in. in steps of 0.1 in. at the other end.

Procedure. The procedure consisted of three steps. The first was to estimate the width of the standard before stimulation (standardization judgment). S stood, blindfolded, between two tables which supported the standard and the scale. He held the 2-in. standard strip between the thumb and fingertips of one hand. With the other hand, he grasped the graduated scale, which was moved by E until S stated that the width of the two objects was equal. Six individual 'readings,' three ascending and three descending, were taken to obtain a mean value. In the second step, 'interpolated stimulation,' S grasped the interpolated stimulus-strip with the hand previously used for holding the standard, and ran his hand up and down the strip for the duration of 1 min. The third step was a repetition of the procedure followed during standardization ('test width-judgment'). Change in apparent width was calculated in reference to the value obtained in the 'standardization judgment.' Each test-session consisted of two phases, with an intervening rest-period of 5 min.

⁴G. S. Klein and David Krech, "Cortical conductivity" in the brain-injured, *J. Personal.*, 21, 1952, 118-148.

In one phase, the small (1-in.) strip was used for interpolated stimulation, in the other the large (4-in.) strip was used.

Ss were tested in two sessions; in one session the right hand held the standard and interpolated stimulus; in the other, the left hand was used. Since preliminary studies showed that handedness has no influence on the after-effect in the normal S, only one hand was tested in the control group; in half of the cases the dominant hand, in the other half the non-dominant hand was used.

RESULTS

Fig. 1 is a graphic representation of results obtained for all three groups: normal control Ss, brain-injured Ss, with sensory impairment, and brain-injured Ss, without sensory impairment. The points represent group means

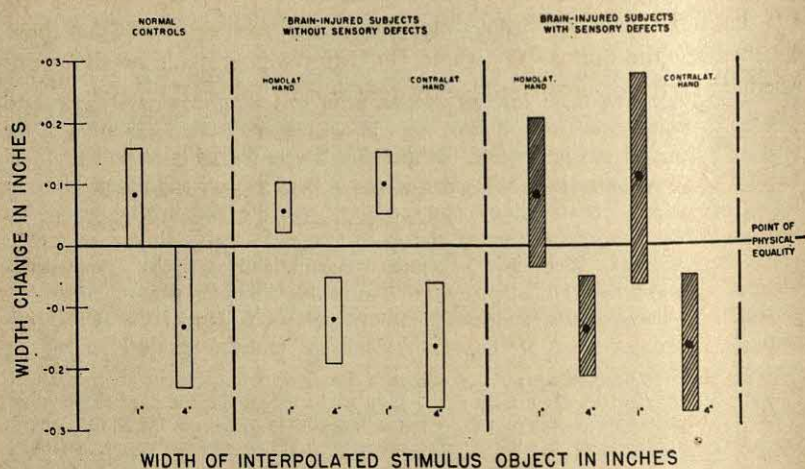


FIG. 1. CHANGE IN APPARENT WIDTH OF STANDARD FOLLOWING INTERPOLATED STIMULATION

(Bars represent group dispersion; points, group means.)

for change in apparent width of the standard after interpolated stimulation. The standard deviation is shown as the length of the bar. Table I presents a summary of the data illustrated in the figure.

In the 20 normal control Ss (Fig. 1a), the standard appeared to increase in width after interpolated stimulation with the small (1-in.) strip, and to decrease in width after stimulation with the large (4-in.) strip. These findings are consistent with those of Köhler and Dinnerstein.⁵ It should be noted, however, that the absolute amount of the effect was much greater after stimulation with the larger strip than after stimulation with the smaller (0.13 as compared with 0.08 in.). Under both conditions, mean

⁵ Köhler and Dinnerstein, *op. cit.*, 209.

after-effects were significant at the 1-% level of confidence as evaluated by the *t*-test.

Results for the brain-injured Ss are shown in Fig. 1b and 1c. These Ss were tested in two sessions that we could assess, separately, the after-effects appearing in the hand opposite to the cerebral lesion (affected hand) and in the hand homolateral to the lesion (unaffected hand). Accordingly, some Ss used the affected hand to hold the standard in the first session, and the opposite (unaffected) hand in the second session; in the other Ss, this procedure was reversed.

The result for brain-injured Ss without sensory impairment can be seen in Fig. 1b. These brain injured Ss obtained after-effects significant at the 1-% level. The effects did not differ in direction or magnitude from those obtained for the normal Ss (Table II). Moreover, it made no difference whether affected or unaffected hands were tested (Table III).

TABLE I
MEAN AFTER-EFFECTS (IN IN.) OBTAINED FROM THE THREE GROUPS OF Ss

		Ss without sensory loss			Ss with sensory loss	
		Controls	homolat. hand	contralat. hand	homolat. hand	contralat. hand
1-in. strip	Mean	+0.08	+0.056	+0.104	+0.07	+0.109
	SD	0.077	0.045	0.047	0.113	0.169
	<i>t</i>	6.153	3.71	6.66	1.85	1.93
	<i>P</i>	<0.01	<0.01	<0.01	0.1	0.1
4-in. strip	Mean	-0.13	-0.124	-0.157	-0.131	-0.165
	SD	0.103	0.070	0.103	0.077	0.106
	<i>t</i>	3.333	5.30	4.65	5.10	4.66
	<i>P</i>	<0.01	<0.01	<0.01	<0.01	<0.01

Fig. 1c shows the results for the 10 Ss who had sustained brain injury and *did* have sensory impairment. Once again the group means are not significantly different from those of the normal controls. The after-effects are significant except where the small (1-in.) strip was used as interpolated stimulus. In those instances, variation between individuals was so great that the mean after-effect failed of significance.⁶

The apparent difference in the performance of 'affected' and 'unaffected' hands in both groups with brain damage did not prove to be statistically significant (Table III). Thus, in all groups, there appeared a marked distortion in apparent size of the standard after interpolated stimulation. For

⁶ The *F*-ratio showed this variability to be significantly greater than that of the other groups.

both brain-injured Ss and normals, there was a greater after-effect following the use of the 4-in. interpolated stimulus than following the use of the 1-in. strip. In no instance did we find a significant difference between normal and brain-injured groups in amount or direction of after-effects. Moreover, no

TABLE II
MEAN DIFFERENCE (IN IN.) BETWEEN THE RESULTS OF THE EXPERIMENTAL
AND THE CONTROL GROUPS

		Ss without sensory loss		Ss with sensory loss	
		homolat. hand	controlat. hand	homolat. hand	controlat. hand
1-in. strip	Mean diff.	0.02	0.02	0.01	0.03
	t	0.732	0.727	0.274	0.644
	P	0.50	0.50	0.80	0.60
4-in. strip	Mean diff.	0.01	0.03	0	0.03
	t	0.267	0.730	—	0.719
	P	0.80	0.50	—	0.50

significant difference was found in any group, depending on the hand used (whether right or left, whether homolateral or contralateral to a brain lesion).

DISCUSSION

The most striking finding in our study is the uniformity in performance of brain-injured and normal Ss. Under comparable conditions of testing, kinesthetic after-effects appeared which were practically identical for all experimental groups. These findings present difficulties for current formula-

TABLE III
MEAN INDIVIDUAL DIFFERENCES (IN IN.) OF AFTER-EFFECTS BETWEEN
AFFECTED AND UNAFFECTED HANDS

		Ss with sensory loss		Ss without sensory loss	
		1-in. strip	4-in. strip	1-in. strip	4-in. strip
1.	Mean diff.	0.04	0.04	0.05	0.05
	SD	0.21	0.12	0.07	0.10
	t	0.65	0.88	2.00	1.34
	P	0.60	0.40	0.10	0.30

tions of the satiation hypothesis. This statement must be qualified, of course, by the validity of the basic assumption involved in our study. We have assumed that fairly gross and localized changes in electrical field activity in the brain might reasonably be expected after brain damage, since such damage involves loss of brain substance with subsequent formation of scar tissue², porencephalic cysts, and meningocerebral adhesions. This as-

sumption is supported by the fact that such cerebral lesions can be shown to produce severe abnormalities in the electrocorticogram and electroencephalogram.

Köhler's conceptions regarding the physiological processes underlying figural after-effects are specific.⁷ They involve the following two assumptions: (a) Rôle of electrotonus; *i.e.* perceptual activity depends on the appearance of direct currents in a corresponding cortical field. (b) Psycho-neural isomorphism; *i.e.*, the distribution of these currents in the field, and their rise and fall in time, are directly related to the shape, localization and duration of the resulting percepts. To support Köhler's hypotheses, evidence is needed on both of the above assumptions. One should be able to demonstrate (a) the reality of the electrotonic changes in the course of perceptual activity, and (b) the appearance of distortions in perception when the underlying neural field is distorted by injury, disease, or by special experimental agents.

In support of the rôle of electrotonus, Köhler and his co-workers have recorded DC potential changes from the occipital scalp which appear related to the passage of visual objects through S's visual field.⁸ Similar potential changes are said to occur in the temporal regions during auditory stimulation. Although the existence of such DC potentials appears to be established, their interpretation has been questioned in view of the many possible sources of artifacts in records of this nature.⁹ Questions have also been raised by experiments designed to disturb the cortical fields by short-circuiting DC potentials. Lashley, Chow, and Semmes inserted gold pins and strips of gold foil in the occipital lobes of two monkeys, and observed no change in the monkeys' subsequent visual performance on a number of discrimination problems.¹⁰ This experiment may not be crucial, however, since it involves two unproved assumptions: (a) that the visual tasks required of the animals were adequate to disclose visual deficits, and (b) that

⁷ Köhler, *Relational determinants in perception*, in L. A. Jeffress (ed.), *Cerebral Mechanisms in Behavior (The Hixon Symposium)*, 1951, 200-243; Köhler and Julia Fishback: The destruction of the Müller-Lyer illusion in repeated trials: I. An examination of two theories, *J. Exper. Psychol.*, 40, 1950, 267-281.

⁸ Köhler and Richard Held, The cortical correlate of pattern vision. *Science*, 110, 1949, 414-419; Köhler, Held, and D. N. O'Connell, An investigation of cortical currents, *Proc. Amer. Phil. Soc.*, 96, 1952, 290-330.

⁹ R. W. Gerard: Discussion, in L. A. Jeffress (ed.), *Cerebral Mechanisms in Behavior (The Hixon Symposium)*, 1951, 236; Donald Lindsley, Discussion, *ibid.*, 241-242.

¹⁰ K. S. Lashley, K. L. Chow, and Josephine Semmes, An examination of the electrical field theory of cerebral integration, *Psychol. Rev.*, 58, 1951, 123-136.

the presence of gold foil or pins interferes sufficiently with the formation of electrotonic fields.

Köhler's second assumption involves the isomorphism between perception and the distribution of DC potentials. Distortions in the underlying cortical field would therefore be expected to produce corresponding distortions in perceptual processes. This assumption derives some degree of support from the generally accepted concept of point-to-point projection; as shown by electro-physiological studies of localized evoked potentials in the post-central and occipital cortex resulting from localized cutaneous or visual stimulation.¹¹ There is also ample evidence of localized visual defects which are correlated with localized lesions in the visual pathways.¹² While this evidence attests to a certain degree of correspondence between specific points in the perceptual field and specific areas in the cerebrum, it does not bear directly on the type of prolonged electrical field changes invoked by Köhler to account for figural after-effects.

A study which applies more directly to Köhler's hypothesis, and which presents negative evidence, was done by Teuber and Bender, who studied apparent motion in Ss who had scotomata due to penetrating lesions of the occipital lobe.¹³ In spite of the anatomical and behavioral evidence of discontinuity in the primary visual cortex, these Ss showed no difference in apparent motion induced across the scotomata as compared with apparent motion in intact areas of the visual field.

Our present study may also be taken as a partial test of Köhler's second assumption, inasmuch as it may be postulated that a brain-injured S with demonstrable sensory loss has suffered direct damage to the cerebral somato-sensory projections. Such lesions could be expected to result in gross disturbances in electrical field activity, and therefore in alterations in figural after-effects. Our Ss included a group with lesions fitting the above criteria. Contrary to what might have been expected on the basis of Köhler's formulations the results of this group did not differ significantly from those of the normal controls (or, indeed from those of the brain-injured Ss without sensory defects).

¹¹ C. N. Woolsey, W. H. Marshall, and Paul Bard, Representation of cutaneous tactile sensibility in the cerebral cortex of the monkey as indicated by evoked potentials, *Johns Hopkins Hospital Bull.*, 70, 1942, 399-441; S. A. Talbot and Marshall, Physiological studies on neural mechanisms of visual localization and discrimination, *Amer. J. Ophthalmol.*, 24, 1941, 1255-1264.

¹² G. Holmes and W. T. Lister, Disturbances of vision from cerebral lesions, with special reference to the cortical representation of the macula, *Brain*, 39, 1916, 34-73.

¹³ H. L. Teuber and M. B. Bender, Perception of apparent movement across acquired scotomata in the visual field. *Amer. Psychol.*, 5, 1950, 271 (abstract).

Contrasting results have been reported by Klein and Krech¹⁴ in a study apparently performed simultaneously with our own.¹⁵ They found an *increase* in the rapidity of onset and magnitude of the kinesthetic figural after-effect in a group of 12 patients with brain tumors and diffuse encephalopathies. Specific data in regard to the sensory status and locus of lesions in these Ss were not reported. In view of the fundamental importance of the concept of point-to-point projection to the Köhler-Wallach hypothesis, data bearing on the locus of lesions are essential in order to adduce positive results such as those of Klein and Krech as critical evidence in support of the 'satiational' theory.¹⁶ It must be demonstrated that patients with known lesions in the primary projection area are the only Ss who show alterations in figural after-effects.

In general, the results of the present study do not bear out predictions based on satiational theory. Our negative results indicate that the phenomenon of figural after-effects in kinesthesia cannot readily be explained on the basis of a theory involving isomorphic projection of sensory impulses and altered patterns of current flow related to changes in 'cortical conductivity.'

SUMMARY AND CONCLUSIONS

Twenty experimental Ss and 20 normal controls were tested for kinesthetic figural after-effects according to Köhler and Dinnerstein's method. The experimental Ss were brain-injured veterans of World War II. Ten of them had evidence of impairment of sensory function, 10 showed no sensory impairment. The Ss were tested for distortions in the apparent width of 2-in. standard strip after interpolated stimulation with strips 1 and 4 in. wide.

There was no significant difference between the performance of the

¹⁴ Klein and Krech, "Cortical conductivity" in the brain injured, *Amer. Psychol.*, 6, 1951, 264 (abstract).

¹⁵ Robert Jaffe, Influence of cerebral trauma on kinesthetic after-effects, *ibid.*, 6, 265 (abstract).

¹⁶ In their report, Klein and Krech specifically reject Köhler's concept of isomorphism and postulate instead the existence of a *generalized* state of conductivity whose basal level and susceptibility to the effects of external stimulation is altered by the presence of 'cortical' lesions. For the most part, their results are not significant in terms of conventional statistics. In spite of this, and in spite of the fact that none of their subjects could be said to have a lesion restricted to the cortex, extensive generalizations as to 'cortical function' are drawn from the trends in their data. The only data which achieved significance were correlations between the amount of after-effects and the degree of general neurological disability. While these results are taken as evidence of a specific physiologic mechanism, they might be equally well accounted for in terms of any number of physiological or behavioral hypothetical constructs.

normal and brain-injured groups. A single exception was found in the brain-injured group with sensory deficit; this group showed after-effects of normal magnitude and direction under all conditions of testing except one, in which variability was so great that the after-effects fell short of being significant. With this one exception, no differences were found between groups, or conditions. In particular, no significant difference in mean results appeared in the performance of the hand contralateral to the cerebral lesion as compared to the hand homolateral to the lesion.

The implications of this study are considered in relation to the satiation hypothesis of Köhler and Wallach. Their hypothesis assigns a primary rôle in the mechanism of figural after-effects to the existence of electrotonic fields in the cerebral cortex. These fields are assumed to be isomorphic with the resulting perceptions. Both these assumptions appear to be incompatible with the negative findings of this study which disclosed essentially normal kinesthetic after-effects in Ss with penetrating cerebral lesions.

A RECONCILIATION OF THE VEG SCALE WITH FECHNER'S LAW

By WALTER C. MICHELS, Bryn Mawr College, and
HARRY HELSON, University of Texas

Attempts to establish psychologically significant ratio-scales have occupied the attention of a considerable number of workers in psychophysics ever since Johnson presented his criticism of 'pseudo-mathematics' in 1935.¹ Fractionation-methods have seemed to offer the best hope of obtaining such scales and have been used in the study of loudness, pitch, warmth, weight, and brightness.² The results obtained when the various sense modalities are tested are alike to the extent that the subjective intensity of the stimulus appears to increase somewhat faster than the logarithm of the physically measured intensity—a result which is in conflict with Fechner's law. It has been pointed out recently that the argument that fractionation-methods lead to ratio-scales is open to logical objection.³ True ratio-scales, as distinct from interval-scales, can be established only if there is a one-to-one correspondence between the intensity of the physical stimulus and the magnitude of the response, *i.e.* if there exists a definite intensity of stimulus which produces a stable zero of response. The fact that both the response to a given physical stimulus and the threshold depend on the adaptation of the organism prohibits us from accepting such a stable zero. It is the purpose of the present discussion to show that the results obtained by fractionation can be explained quantitatively on the assumptions that interval-scales, rather than ratio-scales, are appropriate for psychological purposes and that such scales may be established by techniques used in rating-scales.

One of the best determined ratio-scales, and one which has received considerable discussion, is the *veg* scale for subjective weight, introduced by Harper and Stevens.⁴ Because this scale is very similar to others obtained

* Accepted for publication September 21, 1953.

¹ H. M. Johnson, Pseudo-mathematics in the mental and social sciences, this JOURNAL, 48, 1936, 342-351.

² A review of the concepts involved in methods of fractionation and a good bibliography are given in Chapter 1 of *Handbook of Experimental Psychology*, edited by S. S. Stevens, 1951, 1-49.

³ W. C. Michels and Harry Helson, Man as a meter, *Physics Today*, 6, 1953, (No. 8), 4-7.

⁴ R. S. Harper and S. S. Stevens, A psychological scale of weight and its derivation, this JOURNAL, 61, 1948, 343-351.

by fractionation, and because sufficient data are available on weight-lifting experiments done with rating-scale techniques, the following discussion will be principally concerned with the *veg* scale.

The reformulation of the Fechner law proposed several years ago is based on several assumptions, the ones of special importance to the present treatment being the applicability (within limits) of the Weber law and the supposition that a properly constructed rating-scale will be used by subjects as an interval-scale.⁵ Subsequent experimental tests have offered strong support for the validity of these assumptions, particularly in weight-lifting experiments.⁶ The major prediction of the theory is that the judgment (J) of a stimulus of physical intensity X is given by the expression:

$$J = N + K' \log_e (X/A), \dots\dots\dots [1]$$

where N is defined by the condition that there be $2N + 1$ categories in the rating-scale, assigned values 0, 1, 2 . . . $2N$, K' is a function of N , and A is the adaptation-level. This adaptation-level is a weighted geometric mean of all stimuli, past or present, which affect the judgment. The weighting factors for the various stimuli may be determined from properly designed experiments.

The original theoretical treatment could not be applied to fractionation-experiments, since it dealt only with absolute judgments. It has been shown recently, however, that an expression which is formally identical with Equation [1] applies to comparative judgments.⁷ In this application the neutral category 'equal' is assigned the value N , and the comparative adaptation-level (A) is very strongly influenced by the standard in terms of which the comparative judgments are made. The success of the modified theory in correlating time-order effects with other psychophysical data justifies the extension of rating-scale techniques to comparative judgments.⁸ It therefore seems worth while to investigate the predictions of the theory concerning the results of fractionation-procedures.

In the determination of the *veg* scale, which may be considered typical of such experiments, the procedure was, in part, as follows.⁹ "Each of the

⁵ Michels and Helson, Reformulation of the Fechner law in terms of adaptation-level applied to rating-scale data, this JOURNAL, 62, 1949, 355-368.

⁶ E.g. see M. C. Nash, An experimental test of the Michels-Helson theory of judgment, this JOURNAL, 63, 1950, 214-220.

⁷ Helson, Michels, and Artie Sturgeon, The use of comparative rating-scales for the evaluation of psychophysical data, this JOURNAL, 67, 1954, 321-326.

⁸ Michels and Helson, A quantitative theory of time-order effects, this JOURNAL, 67, 1954, 327-334.

⁹ Harper and Stevens, *op. cit.*, 344.

12 Os made three judgments with each series of weights. . . . The Os were instructed to make successive comparisons between the standard and each comparison weight, and to select that weight 'which feels half as heavy as the standard.' " If these instructions were followed, and if Equation [1] can be applied, it follows that the judgment $J_{1/2}$ of the weight which feels half as heavy as the standard must have a magnitude equal to $1/2 J_1$, where J_1 is the judgment of the standard. Denoting the physical weight of the standard by S and that of the selected weight by X , we have:

$$J_1 = N + K' \log_e (S/A), \dots\dots\dots [2]$$

and

$$J_{1/2} = N + K' \log_e (X/A) = 1/2 J_1. \dots\dots\dots [3]$$

The elimination of J_1 from these two equations shows that

$$\log_e (S/X) = N/2K' + 1/2 \log_e (S/A). \dots\dots\dots [4]$$

Equation [4] may be expected to predict the ratio of the physical weight of the standard to the physical weight that is judged to be 'half as heavy,' provided that the parameters N , K' , and A can be evaluated.

To a first approximation, valid if the standard is not too far different either from the mean of the series stimuli or from other stimuli ordinarily encountered, the comparative adaptation-level has been shown to be approximately equal to the standard.¹⁰ We may therefore expect the general shape of the *veg*-curve vs. the weight-curve to be determined by setting $A = S$, $\log_e (S/A) = 0$, to obtain

$$\log_e (S/X) \approx N/2K'. \dots\dots\dots [4A]$$

In judgments expressed on a rating-scale, $2N + 1$ is the number of categories in the scale. Such a scale is not explicitly introduced into fractionation-experiments, but a close correspondence between these experiments and those using rating-scales may be established. The minimal number of categories which the Os could use, below 'equality,' would be three: 'more than half as heavy,' 'half as heavy,' and 'less than half as heavy.' Hence there would be a minimum of seven categories in a scale which is symmetrical about 'equality.' Since finer discriminations than this would not be needed, it appears that we can take $N = 3$.¹¹

¹⁰ Michels and Helson, *op. cit.*, footnote 8.

¹¹ If this choice is changed to $N > 3$, there will be little effect on the final conclusions, since K' is proportional to $-1/\log_e (1-1/N)$. This quantity, in turn, is very nearly equal to N for $N > 3$, hence the ratio N/K' which occurs in Equations [4] and [4A] is, to a high degree of approximation, independent of N .

The slope, K' , of the J vs. $\log_e X$ curve depends on two factors. Were it not for self-adaptation, *i.e.* adaptation to the stimulus being judged, and if the 'equivalence condition' were satisfied, we might expect the slope to be

$$K = -1/\log_e (1 - 1/N)^{12} \dots\dots\dots [5]$$

To satisfy the equivalence condition, one must present the O s with a series of stimuli ranging over nearly a ten-fold range in weight, as has been shown by Nash.¹³ In the determination of the *veg* scale, the range used, including the standard, was never more than two-fold. Experiments done with such narrow ranges of stimuli have shown that the observed slope K' is very closely equal to $2K$.¹⁴ We may therefore set

$$K' = -2/\log_e (1 - 1/N) = -2/\log_e (2/3) = 4.94. \dots\dots [5A]$$

Substitution of this value and of $N = 3$ into Equation [4A] yields the value $\log_e (S/X) = 0.304$, or $S/X = 1.36$. To the present approximation, this ratio would be expected to be a constant. Harper and Stevens obtained values ranging from 1.27 (20-gm. standard) to 1.55 (1000-gm. standard), with a mean value of 1.43.

Fig. 1 shows a plot of the data of Harper and Stevens (vertical bars). The broken line is the plot of the line $S = 1.36 X$. It will be noticed that the approximate relation of Equation [4A] fits the data within the standard deviations of the judgments, indicated by the vertical bars. There is, however, a systematic deviation of the prediction from the mean judgments, the predicted value being generally too high at the higher values of S . We may now investigate the more exact Equation [4] to see whether the neglected second term on the right improves the prediction.

Before we can apply the correction, we need an estimate of the comparative adaptation-level, A . This quantity is a function of the standard, (S) the geometric mean (M) of the other weights presented to O , and the residual stimulus R . If the logarithms of S , M , and R all lie within a range N/K' of $\log A$, the value is given by:¹⁵

$$\log A = r \log R + m \log M + s \log S, \dots\dots\dots [6]$$

where

$$r + m + s = 1. \dots\dots\dots [6A]$$

¹² Michels and Helson, *op. cit.*, footnote 5, p. 358.

¹³ Nash, *op. cit.*, 217.

¹⁴ Helson, Michels, and Sturgeon, *op. cit.*, p. 14.

¹⁵ Michels and Helson, *op. cit.*, footnote 5, p. 367; footnote 8, p. 329.

If any one of the logarithms of S , M , or R departs from A by more than N/K' , that logarithm must be replaced by $\log A \pm N/K'$ in Equation [6]. The geometrical mean of the weights other than the standard, as listed by Harper and Stevens, may be used for M ; R has been shown to be about 250 gm. in weight-lifting experiments similar to those under consideration;¹⁶ s has recently been measured as 0.85;¹⁷ we may divide the remaining weight-factor ($1 - s$) about equally between r (0.07) and m (0.08) on the basis of the simplest assumption regarding their contributions to the

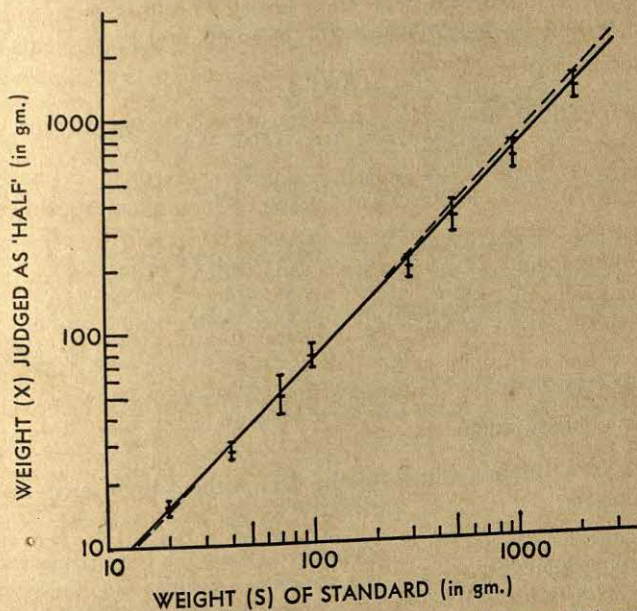


FIG. 1. COMPUTED AND OBSERVED FRACTIONATED DATA FOR WEIGHTS

The vertical bars indicate both the mean values obtained by Harper and Stevens and the standard deviations of the judgments as given by them. The solid line shows the calculated relation between S and X according to Equation [4] and the broken line according to Equation [4A].

comparative adaptation-level. These two contributions are so small compared with that of the standard that the final values will not be affected seriously by errors in these weighting factors.

The calculation of the adaptation-level, the correction term, and the ratio S/X for five different standards is shown in Table I, and the final prediction of the weight that will be judged half as heavy as the standard

¹⁶ Michels and Helson, *op. cit.*, footnote 5, p. 366.

¹⁷ Michels and Helson, *op. cit.*, footnote 8.

is shown by the solid line in Fig. 1. When Equation [4] is used for the prediction, the fit to the experimental data is as good as that obtained by Harper and Stevens with empirical fitting.

The derivation of Equation [4] is quite different in principle from that given by Harper and Stevens,¹⁸ or that published more recently by Armington.¹⁹ These derivations depend on the empirical determination of a function from a log-log plot of the experimental data. This function is then integrated or summed to obtain a new, and more convenient, function. The present treatment starts from principles which have been tested independently of the experiment being discussed, and proceeds to the de-

TABLE I
COMPUTATION BY EQUATION [6] OF THE RATIO OF THE STANDARD (S) TO THE
WEIGHT (X) WHICH APPEARS HALF AS HEAVY

S (gm.)	log S	log M	0.07× log X _r	0.08× log M	0.85× log S	log A	$\frac{1}{2} \log \times$ (S/A)	S/X	
								Calc.	Obs.†
20	3.00	2.72	(0.26)*	0.22	2.55	(3.03)*	-0.02	1.33	1.26
100	4.61	4.41	(0.37)	0.37	3.92	(4.66)	-0.02	1.33	1.30
500	6.21	5.98	0.39	0.50	5.28	6.17	0.02	1.39	1.48
1000	6.90	6.47	(0.44)	0.55	5.86	(6.85)	0.02	1.39	1.55
2000	7.61	7.26	(0.46)	0.61	6.17	(7.24)	0.18	1.63	1.52
							Means	1.41	1.42

* Numbers in parentheses are computed by taking $\log X_r = \log A \pm N/K'$, rather than as $5.52 = \log 250$. All logarithms in this table are to be base e .

† Harper and Stevens' values.

velopment of a formula which fits the data without the use of any arbitrary parameters.

The fact that the data used to establish the *veg* scale can be explained in detail by the use of a theory based on Fechner's law raises a question as to the meaning of ratio-scales. That such scales have genuine operational meaning cannot be doubted, since they are established by a clearly defined and reproducible experimental procedure. Whether their operational basis is broad enough to allow the scales to have a psychological meaning in situations other than that involved in their determination is, however, more doubtful. The apparent flaw in the reasoning underlying their construction is the assumption that a fixed physical stimulus is psychologically equivalent to itself when it is presented under different circumstances. Thus, Harper and Stevens, after defining the subjective weight of 100 gm. as one *veg*, show by fractionation that 72 gm. correspond to 0.5 *veg* if a 100-gm. weight is used as a standard. They then assume that the 72-gm. weight has the same subjective weight (0.5 *veg*) when it is presented as a standard for a new fractionational determination, and thus arrive at the subjective of $\frac{1}{4}$ *veg* for a 52-gm. stimulus. According to Equations [2], [3], and [6], the 100-gm. weight evokes a response of 3.0 on a scale of 6.0 when it is presented as a standard, and the 72-gm. weight evokes a response of 1.5 under these conditions of adaptation. When, however, the latter is

¹⁸ Harper and Stevens, *op. cit.*, 349-351.

¹⁹ J. C. Armington, A note concerning the *veg* scale of apparent weight, this JOURNAL, 66, 1953, 304-306.

presented as a standard, the response to it, except for the negligible correction $\frac{1}{2} \log_e(S/A)$, is 3.0,—the same response as was assigned to the 100-gm. weight before. We may notice that the ratio of the psychological weight of 100 gm. to that of 72 gm. is 2:1, regardless of which is used as standard, and that the same ratio holds for the 72-gm. and 52-gm. weights. Unless, however, the assumption is made that the two responses to the 72-gm. weight are identical, it is not legitimate to multiply these two ratios of 2:1 together to arrive at the conclusion that the ratio of the psychological weights of 100 gm. and 52 gm. is 4:1. Our interpretation is consistent with known facts about adaptation and also with fractionated data; the view that the *veg* scale is a true ratio-scale cannot be reconciled with the well-established concept of adaptation.

Although our approach casts doubt on the validity of the term 'ratio' as applied to the *veg* and similar scales, it does support the contention that *Os* can make judgments of 'half as intense as' with accuracy and reliability. The assumption that such judgments can be made underlies Equation [3] and consequently all of the subsequent treatment—the success of the theory in predicting the experimental results is therefore a direct justification for believing that observers can act in this way.²⁰

Because of the wide applicability of both Equation [1] and of the methods of fractionation, there appears to be no reason why Equation [4] should not apply to the data used to determine the sone and other scales, as well as that considered here. The only apparent difficulty is that parameters such as the residual stimulus and the weighting factors for the standard and series stimuli have not yet been measured for all sense modalities. A study of the data which lead to the 'bril' scale of brightness²¹ has been carried on by one of us and it appears that these data are completely consistent with the reformulated Fechner law.²²

²⁰ After the first part of this study was completed, the authors had the opportunity of seeing a report of some recent unpublished work of J. P. Guilford and H. F. Dingman. The purpose of their experiments was to test fractionation techniques by comparison with constant sum methods. As part of their procedure, they repeated the work of Harper and Stevens except that they offered the *Os* the choice of 17 or 21 weights, varying between 6% of the standard and two or three times the standard. Under these circumstances, they obtained a linear relation between the logarithms of the physical and subjective weights, as did Harper and Stevens, but the slope of the log-log curve is appreciably different from that found in establishing the *veg* scale. Their ratios of S/X vary between 1.67 and 1.92, with a mean value of 1.76.

These recent experiments should correspond more closely to our theory than do the old ones, since the *Os* always worked with a series varying over a range of 10:1 or more. As a result, we may use the theoretical slope K in Equations [4] and [4A] instead of the corrected slope K' . This changes $\log(S/X)$ in the latter equation by a factor of exactly 2.00. We therefore have $\log_e(S/X) = 0.609$ and $S/X = 1.84$. The close agreement between this theoretical value and the ratio found by Guilford and Dingman not only offers additional support for the theory but also shows that apparently discordant experimental results may be reconciled in detail through the use of the reformulated Fechner law.

²¹ Michels, Interpretation of the bril scale of subjective brightness, *J. Opt. Soc. Amer.*, 44, 1954, 70-74.

²² This work has been done as part of a psychophysical program which has been aided by grants from the income of the Carnegie Corporation Endowment given to Bryn Mawr College for the coördination of the sciences.

THE MAGNITUDE OF FIGURAL AFTER-EFFECTS AS A FUNCTION OF THE DURATION OF THE TEST-PERIOD

By JOHN KRAUSKOPF, University of Texas

In a recent discussion, Köhler stated that figural after-effects occasionally appeared to grow during the first few moments of the test-period.¹ Should this observation be borne out by rigorous experimentation, it would not only have important implications for the Köhler-Wallach theory of satiation,² but it would also provide serious evidence against the interpretation of figural after-effects developed by Osgood and Heyer.³ The present experiment was designed to provide a careful test of the reported phenomenon.

In the Köhler-Wallach system, regions of different field-brightness result in regions of different electrical potential in the visual projection-area. It is further hypothesized that direct (figure) currents flow between such regions around the contours. The relative positions of the various contours in the field are supposed to be a function of the 'interaction' of such figure currents. The flow of figure currents is assumed to be a self-limiting process, *i.e.* the passage of these currents alters the medium in which they flow so as to reduce the strength of subsequent currents. According to Köhler, "any direct currents which flow through the nervous system polarize the surface of the cells, and also change their polarizability."⁴

When a contour of a test-figure (*T*) falls in the region of a contour of a previously fixated inspection-figure (*I*), its figure currents encounter an altered cortical medium. If the *T*-contour falls to one side of the *I*-contour, its figure currents will be asymmetrically affected and the contour will appear shifted away from the region of the *I*-contour.

Köhler lays particular emphasis on the polarizability induced by electron-tonus, stating:

One is tempted to say that prolonged presence of an *I*-object lowers the conductivity of its cortical area and adjacent regions. Actually, the change appears to be mainly an increase in polarizability of the tissue in question. . . . If the conductivity of an area is lowered, the intensity of a current which passes through this area is at once decreased. On the other hand, if the same area assumes a high degree of polarizability, a test current will be weakened only as its own flow polarizes the area to the high degree which is now possible. This takes a certain time.⁵

If an increase in polarizability is the dominant result of prolonged fixation, it should then be expected that when exposures of *T*-objects are short, their displacements should be smaller than with longer exposures.

* Accepted for publication November 10, 1953.

¹ Wolfgang Köhler, Relational determination in perception, in L. A. Jeffress (ed.), *Cerebral Mechanisms in Behavior, the Hixon Symposium*, 1951, 200-230.

² Köhler and Hans Wallach, Figural after-effects, *Proc. Amer. Phil. Soc.*, 88, 1944, 269-357.

³ C. E. Osgood and A. W. Heyer, Jr., A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1952, 98-118.

⁴ Köhler, *op. cit.*, 212.

⁵ Köhler, *op. cit.*, 233.

The alternative explanation of figural after-effects developed by Osgood and Heyer is more in accord with orthodox conceptions of neurophysiology. This theory is based, at least in part, on the Marshall and Talbot analysis of the neural mechanisms underlying sensory acuity.⁶ According to Marshall and Talbot, optical factors result in a blurring of the details of the retinal image. Further diffusion of the neural pattern of firing results from the shifting of the retinal image as a function of fixation-nystagmus. The result is that a line is represented in the nervous system as a ridge of activity, normal in shape, rather than an all-or-none copy of a line. Although peaking of the distributions of activity by reciprocal overlap and summative processes is postulated, the cortical pattern of firing is assumed to retain the statistical character of the retinal image.

Fibers of the on-off type are excited by the changes in incident energy caused by shifts of the retinal position of contours due to nystagmus. Osgood and Heyer postulate "that the apparent localization of a contour in subjective visual space coincides with the localization of maximal excitation in area 17."⁷ Further, they postulate that prolonged activity of these on-off fibers results in a state of adaptation, the degree of which is proportional to the rate and duration of firing. In addition, they hold that there is a maximal level of adaptation possible, and that the rate of adaptation is proportional to the difference between the existing level of adaptation and this maximal level. An increase in the level of adaptation results in a proportional reduction of firing.

Figural after-effects are accounted for, according to this model, in the following manner: Prolonged fixation of an *I*-figure results in adaptation in the region of its contours. When a contour of a *T*-figure falls somewhat to one side of the contour of the *I*-figure, the distribution of firing of on-off fibers associated with it is altered as a result of the inhomogeneous pattern of adaptation. The maximum of the distribution, and therefore the apparent location of the contour, is shifted away from the region of the contour of the *I*-figure.

Application of this theory leads to the prediction that after-effects should decrease in magnitude as a function of testing time. If the *T*-figure remains in view for a period of time, adaptation processes due to its presence should occur. Since rate of adaptation is proportional to the difference between the existing level of adaptation and the maximal possible level, the level of adaptation should rise less rapidly in the previously adapted region. In Fig. 1, the right curve shows the pattern of firing which would occur if the *T*-contour were presented in an unaffected region. The left curve indicates the pattern of adaptation which would result from the prior presence of an *I*-contour. The shaded area represents the initial reduction of on-off firing which occurs when the *T*-contour falls in the region of an *I*-contour. Since there is less firing and more adaptation on the left than on the right of the *T*-contour,

⁶ W. H. Marshall and S. A. Talbot, Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity, in H. Klüver (ed.), *Visual Mechanisms, Biological Symposia*, 7, 1942, 117-164.

⁷ The association of the apparent location of contours with maxima of excitation is perhaps the weakest feature of the Osgood and Heyer analysis. Marshall and Talbot (*op. cit.*, 143-149) do not accept this notion in their analysis of irradiation phenomena. Another serious difficulty facing Osgood and Heyer is the specification of a correlate for the apparent thickness of narrow lines (*i.e.* lines so thin that there is only one mode in the distribution of on-off firing).

adaptation takes place faster on the right than on the left, and both level of adaptation and rate of firing on the right and left tend to be equalized. In this way the inhomogeneity of the adaptation-pattern should be reduced and the after-effect should be smaller with longer test-times.

Apparatus. A Gerbrands tachistoscope was modified to permit successive presentation of three stimulus-configurations (a composite of the three configurations is

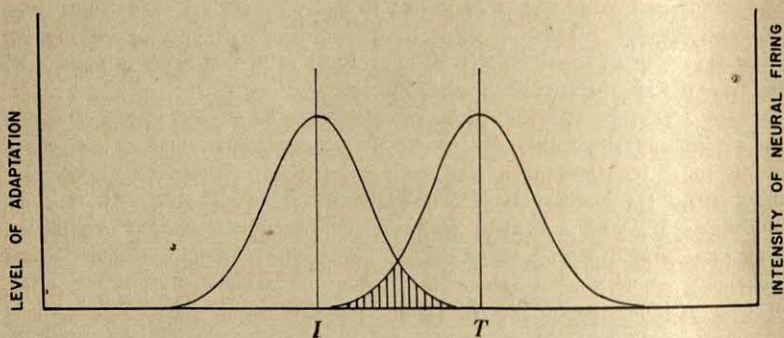


FIG. 1. HYPOTHETICAL DISTRIBUTIONS OF ADAPTATION AND NEURAL FIRING
Left curve-pattern of adaptation associated with the *I*-contour; right curve-pattern of firing associated with *I*-contour falling in an unsatiated region; shaded region—initial reduction in firing when *I*-contour falls in satiated region.

presented in Fig. 2). The configurations were: (1) an *I* figure—a black square ($2^{\circ} 38' \times 2^{\circ} 38'$) and a fixation-point; (2) a luminous fixation-point, and (3) a *T*-figure—two black rectangles ($2^{\circ} 38' \times 1^{\circ} 19'$) and a fixation-point (by appropriate adjustment all three fixation-points appeared in the same place). The first and third configurations were presented by illumination of the front face of the

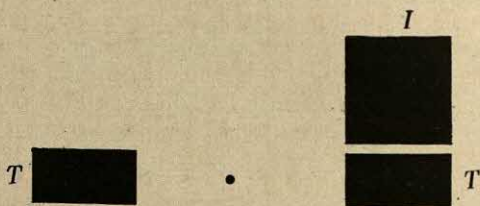


FIG. 2. COMPOSITE OF THE AFTER-EFFECT STUDIED, SHOWING
THE RELATION OF THE *I*- AND *T*-FIGURES

stimulus-figures; the second by shining a light through a hole which served as the fixation-point of the first configuration. The inner edges of the *T*-rectangles and the *I*-square were $2^{\circ} 38'$ from the fixation-point in the lateral direction. The difference in height of the lower edge of the *I*-figure and the upper edge of the right *T*-rectangle was $15.8'$. The left member of the *T*-figure could be moved

vertically by a micrometer control in steps of 0.8'. The presentation of the stimuli was timed by means of a bank of three electronic interval timers of the Hunter-Brown type.⁸

Procedure. The after-effect (the apparent drop of the right rectangle as a function of prior viewing of the square) was measured for five presentation times of the *T*-objects—0.3, 0.6, 0.9, 1.2 and 1.5 sec. After establishing a high level of satiation, method of limits determinations of the setting of the left rectangle of the *T*-figure necessary for it to appear equal in height to the right rectangle were made. Four series (two ascending and two descending) were used to determine the *PSE* for each of the five times. The *T*-objects were presented in the following fashion: After a preparatory signal *E* depressed a key; the *I*-figure was replaced by the fixation-point which was in time replaced by the *T*-figure. The *T*-figure remained on for a short time and was finally replaced by the *I*-figure which remained on until *E* again depressed the key. (In this way, satiation was held relatively constant throughout the session.) To control the reduction in magnitude of after-effects as a function of the time *between* the presentation of the *I*- and *T*-figures, as reported by Hammer, the total time from the disappearance of the *I*-figure and its return was held constant (at 2.1 sec.) for all series.⁹ Before the satiation period, the *PSE* for each test-time was determined for each of the *O*s. For these control settings the preexposure-field contained only the luminous fixation-point. In all other respects the determinations were made in the same fashion as the post-satiation settings.

Five graduate students, whose eyesight was normal or corrected to normal, served as *O*s. One *O* was assigned to each of the five orders of the exposure time determined by a 5 × 5 Latin square. Following the determination of pre-satiation *PSE*s for all exposure-times each *O* viewed the *I*-object for 3 min. (it was expected that a fixation-period of this length would yield maximal satiation). The *O*s then made post-satiation settings for each of the five exposure-times (the same order of exposure-times was used for the post-satiation settings). All the data for each *O* were collected in a single session.

Results. Results of analyses of variance performed on the pre- and post-satiation settings for all *O*s are presented in Tables I and II. For the analyses, the mean of the four settings for each exposure-time for each *O* was used as a score. No significant variance due to exposure-time or trials is evident in the pre-satiation results. The variance due to *O*s is significant beyond the 1-% level of confidence. The variance due to *O*s is also significant beyond the 1-% level for the post-satiation settings. In addition, significant variance attributable to trial (beyond the 5-% level) and to exposure-times (beyond the 1-% level) were found.

No formal test of the significance of the differences between pre- and post-satiation *PSE*s was made since, in the case of every *O*, there was no overlap between the mean pre- and post-satiation *PSE*s.

Since no significant differences attributable to exposure-times or trials was found in the pre-satiation settings, the mean of these determinations for each *O* was taken

⁸ T. A. Hunter and J. S. Brown, A decade-type electronic interval-timer, this JOURNAL, 62, 1949, 570-575.

⁹ E. R. Hammer, Temporal factors in figural after-effects, this JOURNAL, 62, 1949, 337-354.

as the pre-satiation *PSE*. The magnitude of the figural after-effect was computed by subtracting the pre-satiation *PSE* from the mean post-satiation setting for each *O* at each exposure-time. The mean magnitudes (in minutes of arc) for all *O*s are plotted as a function of exposure-time in Fig. 3. A strong inverse relationship between magnitude of after-effect and exposure-time is evident. The smooth curve fitted to the data shows a negatively accelerated form. The relation between magnitude of after-effect and order of presentation presented in Table III is also rather

TABLE I
ANALYSIS OF VARIANCE OF THE LATIN SQUARE DESIGN
FOR THE PRE-SATIATION SETTINGS

Source of variance	Sum of square	df	Mean square	F
Times	3290	4	822.5	.77
Trials	1580	4	395.0	.37
Os	31430	4	7857.5	7.37*
Residual	12800	12	1066.7	
Total	49100			

* Significant beyond the 1% level.

interesting. Although the *I*-figure was presented almost continuously throughout the testing-session, the magnitudes of the after-effects were smaller at the end of the test-session than in the beginning.

Discussion. The results presented show quite clearly that the magnitude of figural after-effects decreases with the length of the testing-period. Thus, it seems that Köhler's contention that the passage of figure currents results in an increase in the polarizability of the tissues is not substantiated. It is, of course, possible that the

TABLE II
ANALYSIS OF VARIANCE OF THE LATIN SQUARE DESIGN
FOR THE POST-SATIATION SETTINGS

Source of variance	Sum of square	df	Mean square	F
Times	54634	4	13658.5	12.96*
Trials	15774	4	3943.5	3.75†
Os	72614	4	18153.5	17.22*
Residual	12652	12	1054.3	
Total	155674	24		

* Significant beyond the 1% level.

† Significant beyond the 5% level.

times used in the present experiment were too long. The curve presented in Fig. 3 is, however, negatively accelerated and there is no indication that the shape of the curve would change if determinations were made for shorter test-times.

The quantitative facts revealed seem to present certain difficulties for the theory presented by Osgood and Heyer. The mean magnitude of the after-effects obtained with a test-period of 0.3 sec. was 12.6'. This value is attenuated by several factors: the interval between the disappearance of the *I*-figure and the appearance of the *T*-figure (1.8 sec. for the 0.3 sec.-exposure), the decrement due to the trials effect, the use of solid rather than outline figures, and the fact that the distance between

the *I*- and *T*-figures was arbitrarily chosen and not necessarily that which would yield the maximal effect. Much larger after-effects than those reported here should be ex-

TABLE III
MEAN MAGNITUDE OF FIGURAL AFTER-EFFECTS FOR
ALL OS ON EACH TRIAL (IN MINUTES)

Trial				
1	2	3	4	5
10.5	10.2	9.1	7.7	9.3

pected (in fact, much larger ones have been observed in this laboratory by use of the present method). Marshall and Talbott¹⁰ give 4' as the mean nystagmic fluctuation, a

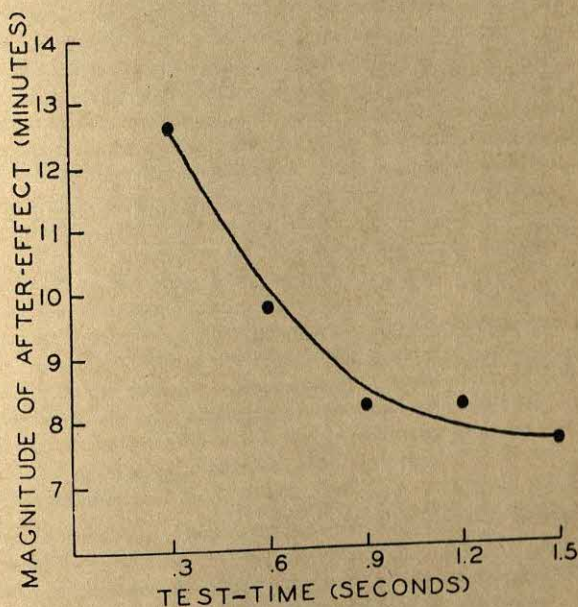


FIG. 3. MEAN MAGNITUDE OF AFTER-EFFECTS AS A FUNCTION OF TEST-TIME
After-effect in minutes of arc; test-time in seconds.

value apparently derived from Adler and Fliegelman.¹¹ The neural mechanisms are supposed to peak the distributions of firing. The sigmas of the distribution of

¹⁰ Marshall and Talbot, *op. cit.*, 144.

¹¹ F. H. Adler and M. Fliegelman, Influence of fixation on the visual acuity, *Arch. Ophthalm.*, 12, 1934, 475-483.

cortical activity must thus be expected to be considerably smaller than 4'. It does not seem possible that after-effects of the magnitude reported here can be derived from such restricted distributions.

Certain methodological conclusions can be drawn from these data. The results show that after-effects can be maximized with short-test exposures. It is also apparent that reliability of measurement can be greatly improved if this variable is properly controlled.

SUMMARY

Figural after-effects have been found to diminish in magnitude as a function of the length of the test-period. The results have been evaluated in the light of the Köhler-Wallach model and the Osgood and Heyer interpretation of figural after-effects.

SYSTEM IN THE RELATIONSHIP OF AFFECTIVE VALUE TO FREQUENCY AND INTENSITY OF AUDITORY STIMULI

By J. P. GUILFORD, University of Southern California

In 1941, Singer and Young published a report of a number of studies of judgments of affective responses to various types of stimuli.¹ Among these studies was an experiment on the relation of affective value to the properties of auditory-stimuli. Stimuli giving relatively pure tones were varied systematically over wide ranges of both frequency and intensity.

Seldom does an investigator envisage his problem so broadly or take the pains to explore his region of investigation so intensively. In the treatment of their data, however, Singer and Young did not take full advantage of the possibilities for showing the continuity and order with respect to the functional dependence of affective values upon the two major stimulus-variables. By combining tones into groups of two intensity-levels while showing the average 'affective values' in relation to frequency and by combining the same tones into groups of two frequency-levels while showing the average 'affective values' in relation to intensity, they were able to draw some very rough conclusions.² These conclusions were that low-pitched tones and soft tones are preferred to high-pitched tones and loud tones.

In a somewhat parallel study on the relation of affective value to hue, tint, and chroma of colors, the writer was able to show considerable continuity and system.³ It is the purpose of this short article to treat the Singer-Young data in a manner analogous to that for color preferences, to see whether we can derive a picture of the shape of the regression of affective value as a function of frequency when intensity is held constant at different levels and of the shape of the regression of affective value as a function of intensity when frequency is held constant at different levels. If there is sufficient clarity of continuity in these relationships, we can also determine what combinations of intensity and frequency will most probably give rise to an average affective reaction of indifference or of any other selected value, pleasant or unpleasant, on the affective scale.

Experimental conditions. Certain experimental conditions are pertinent to understanding and interpretation of the findings to be reported here. Singer and Young report that the tones were produced by means of a beat-frequency oscillator in connection with a dynamic speaker. The selected frequencies varied from 100 through 1000 in steps of 100~, and also included 1200, 1400, 1600, 2000, 2400, 2800,

* Accepted for publication September 6, 1953.

¹ W. B. Singer and P. T. Young, Studies in affective reaction: II. Dependence of affective ratings upon the stimulus-situation, *J. Gen. Psychol.*, 24, 1941, 303-325.

² The term is in quotation marks here for the reason that the authors used percentages as measures of pleasantness and unpleasantness and hence they were dealing with ordinal values. Later the same term will be applied to interval-scale values on the same continuum.

³ J. P. Guilford, System in color preferences, *J. Soc. Motion Pic. Eng.*, 52, 1949, 197-210.

3000, 3200, and 4000~, or 18 in all. The tones of the 5 highest frequencies were each presented at 8 levels of intensity, calibrated in odd-numbered steps from 7 through 21 and all other tones were presented at 6 levels, 9 through 19. There were thus 118 different tones altogether. Every tone was sounded for a duration of 5 to 8 sec. The observers (*O*s) sat in an arc drawn 10 ft. from the speaker. Thirty-nine *O*s judged each tone for degree of pleasantness or unpleasantness 4 times (156 judgments for each tone), the different tones being presented randomly in counter-balanced order. A rating scale was used but the information extracted from the ratings and reported in the Singer-Young article was in terms of percentage of *P* (pleasant) judgments, where the number of 'indifferent' judgments had been divided equally between *P* and *U*.

Quantitative treatment of the data. It is well recognized that percentages are not measurements on an interval scale. Since they are ordinal values only we cannot employ them in an attempt to achieve a picture of the functional relationships between affective value and the stimulus-variables. For this purpose we need interval-scale values. If we assume that the dispersion of affective reactions to any stimulus is Gaussian and if we assume also that the amount of dispersion on the affective scale is the same for all of the stimuli in the experiment, we have a logical basis for converting the given percentages of *P* judgments into corresponding deviate or *z* values on the base of the normal curve. The writer took these steps with the Singer-Young data, then plotted the resulting affective values as a function of frequency (intensity constant) and also of intensity (frequency constant).

There were numerous irregularities in the data, as one should expect from similar experiences. Although the *N* for the percentages was 156, since there were only 39 *O*s, the number of degrees of freedom for each datum is somewhere between those two figures, depending upon how self-consistent the *O*s were. In spite of the irregularities, there was evident system in both plots, and the drawing of each (smoothed) curve could be facilitated by the trends in neighboring curves. Also, the drawing of the curves in the one plot (intensities constant) could be facilitated by the drawing of the curves in the other plot (frequencies constant). That is, for any pair of stimulus-values for intensity (*I*) and frequency (*F*) there had to be an identical affective value in both plots. After considerable smoothing, making use of this cross-information, the resulting fitted functions are as shown in Figs. 1 and 2.

Fig. 1 shows the regressions of affective value on sound frequency at the eight different levels of intensity. Since there was a scarcity of data for *I*-levels of 7 and 21, those curves are in dotted lines. They were completed by introducing extrapolations. It is obvious, however, that at any frequency-level the greater the sound-intensity the less pleasant the tone, and for all intensity-levels the higher the frequency the less pleasant the tone. In the relation to frequency, however, the decline in affective value with frequency is of decreasing deceleration, with little change in affective value above 1500~ for the louder tones, and above 2000~ for the softer tones.

Fig. 2 shows the regressions of affective value on intensity-values for different frequency-levels, only 11 of the 18 experimental frequency-levels being shown. The decline is slowest among the softer tones with generally increased rate of change among tones of moderate intensity and with slower changes for the higher-pitched tones among the louder levels. Thus most of the regressions are slightly sigmoid, exceptions occurring at frequencies of 100 to 300~.

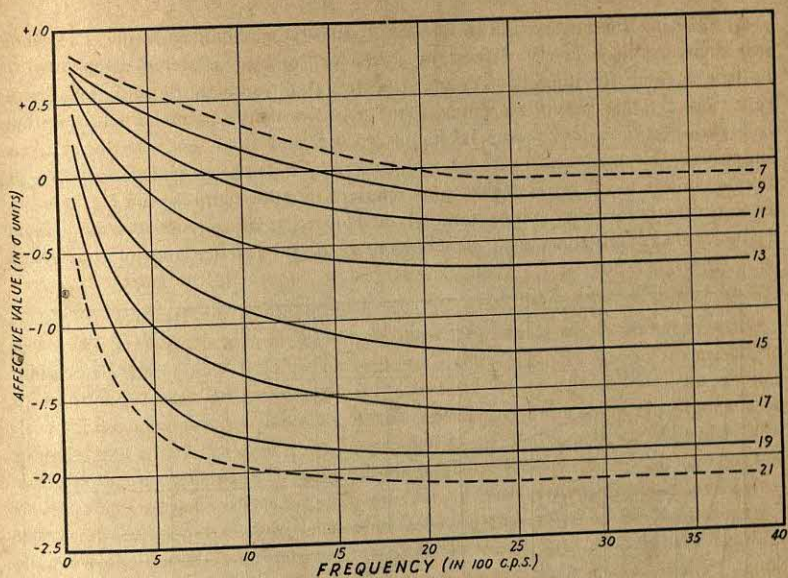


FIG. 1. AFFECTIVE VALUE OF PURE TONES AS A FUNCTION OF STIMULUS-FREQUENCY AT EIGHT INTENSITY-LEVELS
(Data adapted from Singer and Young)

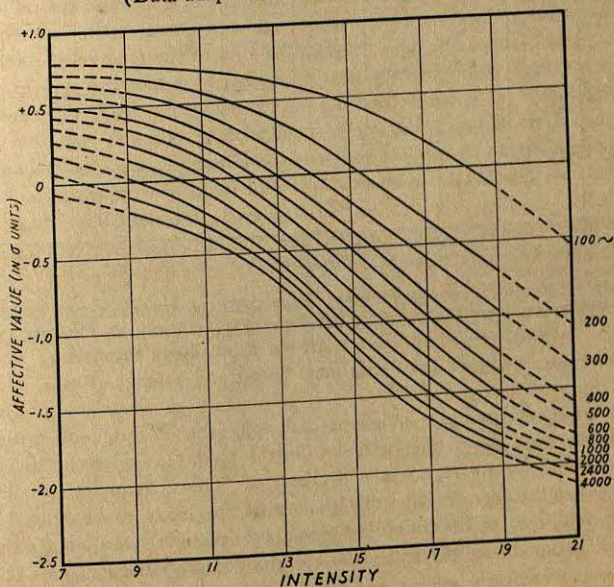


FIG. 2. AFFECTIVE VALUE OF PURE TONES AS A FUNCTION OF STIMULUS-INTENSITY AT ELEVEN FREQUENCY-LEVELS
(Data from Singer and Young)

By utilizing the curves in Figs. 1 and 2, it is possible to construct a chart of isohedons. An isohedon is defined as a line of constant affective value.⁴ Fig. 3 shows the result for the Singer-Young data. In a field representing all combinations of *F*- and *I*-values within the range of the experiment, the points at which certain constant affective values should be found are indicated. First note the locus of the indifference level, at which affective value is zero. It is the second curve from the bottom in Fig. 3. It starts with a tone whose *F* is 100~ and whose *I* is 19, and ends (by extrapolation) with a tone whose *F* is approximately 2600~ and whose *I* is near 7. For all tones with combinations of *F* and *I* above and to the right of

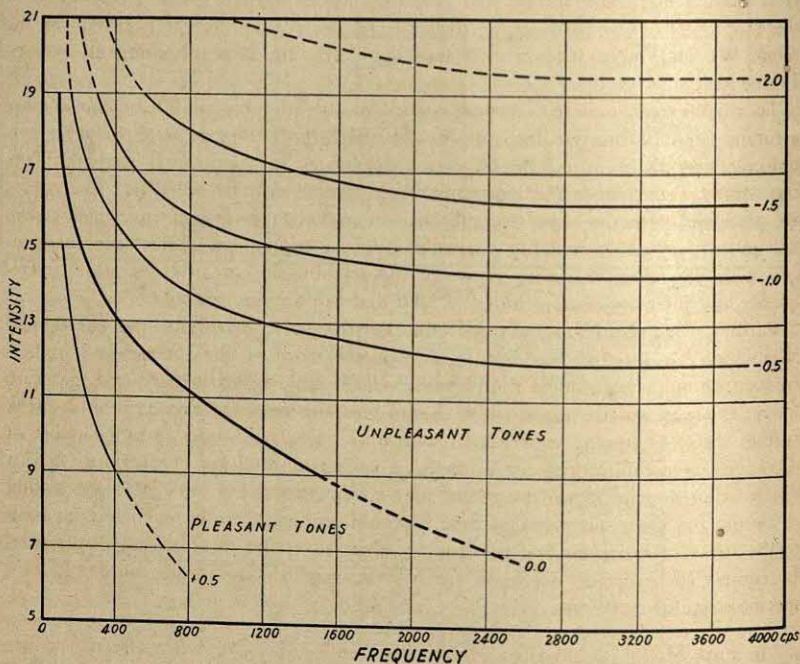


FIG. 3. AN ISOHEDON CHART, SHOWING THE LINES OF CONSTANT AFFECTIVE-VALUE LEVELS AT ONE-HALF-SIGMA INTERVALS FOR VARIOUS COMBINATIONS OF FREQUENCY AND INTENSITY OF PURE-TONE STIMULI
(Based upon data from Singer and Young)

this curve we would predict a judgment of 'unpleasant.' For all tones with *F* and *I* below and to the left of this curve we would predict a judgment of 'pleasant.' Thus, of all the surface represented by the chart, only a relatively small segment in the lower-left corner includes the tones that are likely to be judged pleasant under the conditions of the Singer-Young experiment. For tones among the higher frequencies, variations in affective value are produced almost entirely by changes in

⁴ Guilford, A study in psychodynamics, *Psychometrika*, 4, 1939, 1-21.

intensity. For tones of low frequency, changes in affective value are produced more by changes in frequency.⁵

Some qualifications. It must be remembered that the stimulus-variables used in these illustrations are derived from the calibrations of the sound-producing apparatus; they do not fully represent stimulus-conditions at the ear of the *O*. We are assured by Singer and Young that the frequencies were well controlled and were checked. As the frequencies were changed, however, the intensity would also change somewhat, even with the same intensity-setting of the apparatus, as the authors point out. How much this effect modified the intensity-level at the ear of the *O* is hard to say. The authors reported that checks showed the intensity-levels of the tones varied less than six decibels for *O*s seated at different positions in the room. We may more appropriately take any *I* reading as representing an average *I* value for the *O*.

To what extent a more complete control of the stimulus-conditions and a more accurate quantification of intensity would affect the curves is hard to say. It is unlikely that there would be important departures in the general shapes of the regressions. Some minor readjustments would undoubtedly be called for. The curves are presented with the hope that they will be a challenge to some investigator who will go to the trouble of doing a more painstaking experiment of the same character and with the belief that they represent the relationships of affective value to frequency and intensity of pure tones, at least to a first approximation.

Further work should also consider the probable value of varying the duration of stimulation. It may be that with shorter times more of the intense and higher-frequency tones would be on the pleasant side. It can possibly be predicted that with longer exposures of the tones, the region of pleasant tones (of Fig. 3) would shrink further. From beginning experiments with pure tones, one could go in a number of directions of complication, for example, a systematic addition of overtones, and a presentation of tone sequences and of tone combinations. For those who are willing to devote the great quantities of time it would require, there is practically no limit to the stimulus-variations and the study of their effects on pleasure and displeasure. It requires such studies to achieve the laws from which predictions can be made in specific stimulus-situations.

⁵ In spite of the irregularities in the data mentioned above, when affective values were predicted for the 118 tones on the basis of the isohedon chart in Fig. 3, they were found to be correlated with the experimental *z* values to the extent of +0.94.

A STUDY OF THE 'COMMON CHEMICAL SENSE'

By MARGARET HUBBARD JONES, University of California, Los Angeles

An examination of the literature on the chemical senses reveals a confused state of affairs. In addition to smell and taste, a 'common chemical sense' is accepted by some writers without much tangible evidence and summarily rejected by others without much better evidence.

The term 'common chemical sense' appears to have originated with Parker in 1912.¹ He used it to describe sensitivity to chemicals on the bodies of fishes, and believed the receptors to be free nerve endings. Parker and Stabler in 1913 reported thresholds of irritation for ethyl alcohol in man for non-gustatory regions of the mouth.² They compared these thresholds with those for taste and smell, but present no evidence that the irritation is different from that of ordinary trigeminal pain.

Crozier stated that on the cocaineized frog's foot he was able to obtain differential responses to formic acid and to pinching, with the acid giving good responses with reaction-times of 10 to 15 sec. after the foot had ceased to respond to pinching (touch plus pain?).³ He also tried to separate sensitivity to mechanical and to chemical stimuli in his own mouth by the use of cocaine. No description of method and no data are presented, but Crozier concludes: "I am certain that the two sets of receptors are not only qualitatively distinct as regards the sensations with which they are connected, but also may be separated by the use of cocaine."⁴

Moncrieff devotes a chapter to the topic, and claims, citing Parker and Crozier, that the sense is distinct from those of taste, odor, touch, and pain.⁵ Morgan and Stellar follow Moncrieff.⁶ Geldard ignores the topic completely.⁷ Woodworth states specifically that ordinary pain is responsible for the reaction to chemical irritants in the mucous membranes but cites no evidence for his statement.⁸ Pfaffman is not satisfied with Crozier's evidence. He attributes Crozier's results with the frog's foot to spatial summation and concludes: "Pain and the common chemical sensitivity appear therefore to be mediated by the same nerve endings."⁹

The evidence for the 'common chemical sense' as being distinct from ordinary pain is inconclusive, at least as far as man is concerned. The present study was

* Accepted for publication September 24, 1953.

¹ G. H. Parker, The relation of smell, taste, and the common chemical sense in vertebrates, *J. Acad. Sci., Phila.*, Ser. 2, 15, 1912, 221-234.

² Parker, and E. M. Stabler, On certain distinctions between taste and smell, *Amer. J. Physiol.*, 32, 1913, 230-240.

³ W. J. Crozier, Regarding the existence of the 'common chemical sense' in vertebrates, *J. Comp. Neurol.*, 26, 1916, 1-8.

⁴ *Ibid.*, 6.

⁵ R. W. Moncrieff, *The Chemical Senses*, 1946, 127-146.

⁶ C. T. Morgan and E. Stellar, *Physiological Psychology*, 2nd ed., 1950, 108-109.

⁷ F. A. Geldard, *The Human Senses*, 1953, 1-365.

⁸ R. S. Woodworth, *Experimental Psychology*, 1938, 477.

⁹ Carl Pfaffman, Taste and smell, S. S. Stevens (Ed.), *Handbook of Experimental Psychology*, 1951, 1143-1145.

undertaken therefore to obtain evidence that would be conclusive—either for or against the acceptance of this 'sense.'

Apparatus. Three kinds of stimuli were used to arouse pain: (1) chemical; (2) mechanical; and (3) electrical.

(1) *Chemical.* Solutions of piperidine, the active ingredient of pepper, were prepared with distilled water to give a series of 11 dilutions which extended well above and below 'average sensitivity.'¹⁰ The dilutions were made volumetrically, but these figures have been translated into molar terms (see Table I). The strongest solution of this series produced a significant chemical burn. One drop (0.038 ml.) of solution was applied to the experimental area, allowed to remain for 5 sec., and then the area was rinsed thoroughly. If pain did not result, the next stronger solution in the series was employed, and so on until pain was experienced.

(2) *Mechanical.* A mechanical algesimeter was constructed, utilizing a needle loaded by a weight balanced on a knife-edge.¹¹ The load on the needle was varied by shifting the position of a rider along the lever. The algesimeter was first calibrated on a balance in gram-steps; then quarter-gram steps were laid off metrically. The needle was sharpened to reduce the concomitant stimulation of pressure. In operation the needle was locked in the up position, then the apparatus was brought to within a millimeter of the surface to be stimulated and the needle lowered

TABLE I

SOLUTIONS OF PIPERIDINE USED IN DETERMINING THE THRESHOLD OF PAIN

Solu- tion no.	Ml. of piperidine per 35 ml. sol.	Molar conc.	Solu- tion no.	Ml. of piperidine per 35 ml. sol.	Molar conc.	Solu- tion no.	Ml. of piperidine per 35 ml. sol.	Molar conc.
1	.300	.08628	5	.425	.12199	8	.550	.15799
2	.350	.10057	6	.450	.12914	9	.600	.17256
3	.375	.10771	7	.500	.14371	10	.650	.18656
4	.400	.11485				11	.700	.20085

slowly to the point where the full weight rested on the skin. It is of critical importance that the needle be lowered so slowly that acceleration is a negligible factor otherwise the weight bears little relation to the threshold.

(3) *Electrical.* A break-shock from an inductorium was utilized as the electrical stimulus. The active electrode was a double strand of tantalum wire; the inactive electrode was wet. The thresholds are reported in arbitrary units, there being no reason to attempt to translate them into volts.

¹⁰ Some difficulty was encountered in choosing an appropriate chemical stimulus. Formic acid was tried, since it seemed to be what Crozier had used, but it was so sour and so distracting—even when used on buccal areas lacking taste buds (because of its spread when S rinsed his mouth)—that threshold experiences of pain could not be attended to. The threshold for irritation, furthermore, was so high that accurate measurements, it was feared, would not be possible. Among the numerous chemicals tried, piperidine seemed to be the best for present purposes. We used Eastman's practical grade.

¹¹ This algesimeter is similar to one described by M. H. SeEVERS, J. H. BENNETT, H. W. POHLE, and E. W. REINHARDY, The analgesia produced by nitrous oxide, ethylene, and cyclopropane in the normal human subject, *J. Pharm. & Exper. Therapy*, 59, 1937, 291-300.

Subjects. Seven Ss (graduate students in psychology or members of the staff) were used. Some of them were highly practiced in observation, others were quite naïve.

Procedure. S was told the general nature of the study and the sequence of experiments to be followed. The mucous membrane of the inside of the lower lip was used as the experimental area. Three thresholds of pain were determined by the use of electrical stimulation, three by mechanical stimulation, and one by stimulation with piperidine. All series in the experiment were ascending because of the danger of injury when supra-threshold stimuli are used. After the normal thresholds were obtained, a solution of cocaine hydrochloride was applied to the experimental area, and the thresholds were again determined with the various stimuli. Only one determination of threshold was made for piperidine at a given time because,

TABLE II
AVERAGE THRESHOLD OF PAIN (INSIDE LOWER LIP) BEFORE AND AFTER COCAINIZATION
Stimuli

S	Mechanical		Electrical		Chemical	
	Before	After	Before	After	Before	After
1*	2.80	4.25	6.50	6.00	.115	.144
2	1.83	2.08	7.42	7.50	.144	.201
		3.25†				
3	1.93	3.50	6.75	7.00	.158	.201†
4	2.75	4.58	7.83	7.58	.158	.144
		5.00††§				
5	2.42	5.33†§	7.50	7.50	.144	.158
6	2.25	4.00†	7.33	7.43	.172	.201†
7	1.83	4.83†	7.83	8.33	.144	.186

* This S had a 4-% solution of cocain. His post-cocain thresholds are based on only one series with each kind of stimulus.

† For two Ss, three additional thresholds of pain were obtained for mechanical stimuli after the electrical because the threshold appeared to be steadily rising.

‡ No threshold obtainable in at least one of these trials.

§ On one spot a stimulus of 8.0 grm. was used without reaching threshold. This appeared to have made holes in the epithelium, enabling the chemical stimulus to penetrate to unanesthetized areas.

whenever threshold was reached, the burning or stinging pain persisted for some time, making further work impossible.

Preliminary trials were run using a 4-% solution of cocain which was painted on the experimental area with a camel's hair brush—a procedure used in many early studies on cocainizing taste. When this failed to produce any measurable anesthesia, small cotton pads soaked in the 4-% solution were held against the area for various lengths of time up to 5 min. Only one S showed any effect of anesthetization with this procedure. For the other Ss, the cotton pads were soaked in a 10-% solution. These pads, changed every 2½ min., were applied for 10 min. This procedure produced satisfactory anesthetization, but led to the further difficulty that some of the thresholds were too high to be measured.

Results. The results are presented in Table II. The data are too few for statistical

treatment, hence reliance on trends is necessary. It is apparent that in all cases mechanical pain showed a large increase in threshold after cocaineization. The variability in pre-cocain thresholds was small compared to this rise. The threshold for break-shock was unchanged by cocaineization of the mucous membrane. This is probably due to the fact that there are pain endings at all depths. Since the electrical stimulus penetrated the tissues, it was able to stimulate the endings lying deeper than the cocaine penetrated. Indeed, several Ss remarked during the post-cocain trials that the shock was felt on the *outside* surface of the lip, at the inactive electrode.

Thresholds for piperidine rose after cocaineization in most cases. For Ss 1, 2, and 3 these thresholds showed large increases. For Ss 4 and 5 the thresholds did not rise, but with these two Ss the mechanical stimulus used in an attempt to reach threshold was so intense that it was believed to have made holes in the epithelium, through which the piperidine was able to penetrate to uncocainized levels and to stimulate the receptors there at its normal concentration. For Ss 6 and 7 this possibility was avoided by limiting the mechanical stimulus to twice normal threshold. In these two cases the threshold for piperidine increased as with the first three Ss.

The thresholds of pain aroused by piperidine were peculiar in three ways. In the first place, they seemed unlike normal sensory thresholds because of their definiteness; *i.e.* the sting was either insistently present or there was nothing. Secondly, there was typically a decided time-lag after the application of the 'threshold' solution before the sting became evident, and once aroused, it persisted for some seconds despite repeated rinsing of the area. Finally, every S received chemical burns which were evident for several days, even on spots where the threshold was never exceeded.

Conclusions. In the majority of cases the rise in the threshold for pain is the same, whether obtained with chemical or mechanical stimuli. This suggests that the same mechanism of pain is involved in both instances. There is no evidence for differential effect in those cases where there is no question of faulty procedure.

The peculiar nature of the thresholds obtained with piperidine implies that we are dealing not with a true sensory threshold but with a threshold for actual tissue damage, which results in the subsequent stimulation of pain via the ordinary pain system.

No evidence was uncovered which would lend support to a theory of a separate 'common chemical sense' as distinct from cutaneous pain. The thresholds obtained with chemical and mechanical stimuli do not behave differently under the influence of cocaine. It is probable that chemical stimuli are effective by virtue of destruction of tissue and the ordinary receptors and fibers of pain.

THE DURATIONAL THRESHOLD OF VISUAL RECOGNITION AS A FUNCTION OF WORD-FREQUENCY

By PATRICIA KING-ELLISON and JAMES J. JENKINS,
University of Minnesota

Word-frequency as determined by counts of selected groups of texts has been shown to be an important variable in determining the durational threshold for the recognition of isolated words presented tachistoscopically. Howes has shown that even when legibility of single letters, letter pairs, etc., is not controlled, correlations on the order of -0.65 or -0.75 are obtained for the durational threshold and the logarithm of the frequency of words as given in the Thorndike-Lorge tables.² These results have been of great significance in the criticism of the experiments which have been presented as demonstrating the effect of values and needs on perception.³

Solomon has attempted to examine this perception-time and frequency-relationship in a more rigorous setting in which word-frequency could be directly controlled instead of being estimated from relatively crude methods of counting words. This experiment, which is reported by Solomon and Postman, demonstrated an extremely high relationship (-0.96) between the two variables.⁴

In view of the importance of this relationship to research in the area of perception and verbal behavior, the present experiment was performed to see if the relationship was stable under slightly different conditions.

Method. The experimental procedure was identical with Solomon's except as noted below. Ten experimental words were selected. In this study the words were five-letter paralogues (taken from the portion of Dunlap's list reprinted in Woodworth⁵) in place of the seven-letter Turkish words used by Solomon. A pack of 100 cards, each bearing a paralog, was made up for every S. In each pack two of the paralogues appeared 25 times; two appeared 10 times; two appeared 5 times; two appeared twice and two appeared once, making a total of 10 words and 86 cards. Fourteen cards with dummy words were added to make the total of 100.

* Accepted for publication December 16, 1953. This study is part of a series on verbal behavior being conducted at the University of Minnesota under Contract No. N8 onr-66216 with the Office of Naval Research.

¹ D. H. Howes, *The Definition and Measurement of Word Probability*, Ph.D. Thesis, Harvard University, 1950.

² E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, Teachers College, Columbia University, 1944.

³ See Howes and R. L. Solomon, Visual duration threshold as a function of word probability, *J. Exper. Psychol.*, 41, 1951, 401-410; Solomon and Howes, Word frequency, personal values and visual duration thresholds, *Psychol. Rev.*, 58, 1951, 256-270; and Howes, The intelligibility of spoken messages, this JOURNAL, 65, 1952, 460-465.

⁴ Solomon and Leo Postman, Frequency of usage as a determinant of recognition thresholds for words, *J. Exper. Psychol.*, 43, 1952, 195-201.

⁵ R. S. Woodworth, *Experimental Psychology*, 1938, 15.

Every paralog selected was used at each frequency-level in a counterbalanced design. The order of presentation within a deck was random.

Fifteen students in an introductory psychology class were used as Ss. Every S was told that the paralogs were Turkish words and was asked to spell and pronounce the words on each card in his deck. Following this he was left alone to read unrelated material for 20 min. At the end of this period he was taken to a different room in which the paralogs were exposed to him tachistoscopically. Four dummy paralogs were used for practice and then the 10 experimental paralogs were given in a predetermined sequence. Because of the limitations of time it was not possible to interpose other paralogs in the test-sequence as Solomon had done. While it might appear that this procedure would be transparent and lead to guessing, it did not, as the analysis of the data indicates, affect our results.

Tachistoscopic exposures of each word began at 30 m.sec. Successive exposures increased in steps of 10–20 m.sec. up to 200 m.sec.; in steps of 20–30 m.sec. up to 450 m.sec. and in steps of 50 m.sec. beyond that point. This follows Howes' procedure. The durational threshold was the exposure-time at which the S first correctly identified the test-word. On both the cards and the tachistoscopic material the words were typed in capital letters.

At the end of the experiment, the Ss were allowed to ask questions and to comment. They appeared to have little conception of the aims of the experiment but were concerned as to whether they were giving a good performance.

Results. The data were analyzed to obtain the mean exposure-times for each frequency of presentation in the deck of 100 cards. The data are presented in Table I. The correlation between the mean exposure-time and the logarithm of the frequency of presentation was found to be -0.99 which confirms the results of Solomon and Howes.

Correlations for individual Ss are given in Table II. Here each frequency-class was represented by only two words. Counterbalancing to eliminate the effects of the legibility of the individual words, of the order of the words, etc., was not possible. The median individual correlation was found to be -0.43 which indicated a rather marked effect of frequency even when many other variables were uncontrolled and the single case was considered.

The possibility that the relationship rests upon S's awareness of the experimental aims or upon the elimination of his alternative responses as the test-words are run may be examined by a consideration (1) of the data for the first test-word presented, (2) of a comparison of the recognitive thresholds required for the first and last words, and (3) of the results when they are placed in an information-theory frame of reference.

The average recognitive time of the first word in each of the sets of trials may be grouped by frequency-class. This provides three entries in each class. The average recognitive thresholds are given in Table III. It may be seen that these averages approximate the over-all averages as closely as could be expected with this small number of cases.

A comparison of the first and last paralogs in the test-series shows that the last were recognized, on the average, only 30 m.sec. faster than the first. Since Howes finds the effect of practice in similar experiments up to 60 trials, it seems

quite likely that our difference may be due to that effect. There does not appear to be any such great differences as might arise if the Ss were eliminating alternatives.

A conversion of these data into terms of information-theory is enlightening. If the perceptual situation is considered as a communicational channel through which symbols are being transmitted to S as a destination, the known prior frequency of these symbols may be tested as an index of the amount of 'information' they

TABLE I

MEAN RECOGNITIVE THRESHOLD (IN M. SEC.) FOR EVERY FREQUENCY OF APPEARANCE

Frequency of appearance	25	10	5	2	1
Recognitive threshold	168	203	221	271	289

convey. A highly frequent symbol (one which has a high probability of occurring) conveys very little information, that is to say, it is highly predictable. An infrequent symbol conversely carries a great deal of information.⁶ This may be expressed as the number of binary choices necessary to select the symbol by the formula $H = -\log_2 p$ where p is the probability of the symbol. When this is done we may represent each frequency-class by a figure showing the amount of 'information'

TABLE II

CORRELATIONS OF INDIVIDUAL Ss BETWEEN MEAN RECOGNITIVE THRESHOLD AND LOGARITHM OF FREQUENCY OF APPEARANCE

+ .30	-.31	-.41	-.49	-.66
-.14	-.39	-.43	-.53	-.71
-.16	-.40	-.44	-.66	-.78

which is attached to a member of that class. These data are given in Table IV. Since this manner of expressing the data is a simple transformation of the first measure used, the correlation between 'information' and threshold is the same as the first correlation with the sign changed. (+0.99) This may now be interpreted as follows: As the amount of information per symbol increases, the durational threshold increases linearly. Since this is the kind of relationship we should

TABLE III

MEAN RECOGNITIVE THRESHOLD (IN M. SEC.) OF THE FIRST TEST-WORD FOR EVERY FREQUENCY OF APPEARANCE

Frequency of appearance	25	10	5	2	1
Recognitive threshold	160	187	273	270	287

expect to find in a communication situation in which the probability of a given symbol occurring is the same on every occasion, it also provides support for the argument that elimination of alternatives did not play an important rôle in this experiment.

The data may be expressed by an equation in the form $t = c + kH$ where t is the durational threshold in milliseconds, c is a constant expressing the general time lag in perception (a sort of 'deadtime'), k is the increment in recognitive

⁶ C. E. Shannon and Warren Weaver, *The Mathematical Theory of Communication*. 1949, 1-117.

time for⁶ each additional unit of 'information,' and H is the 'information' in a given symbol expressed in bits. In this experiment c equals 128 m.sec. and k equals 23 m.sec., per bit. As an additional binary choice is necessary to 'select' a given symbol, the cognitive time for that symbol increases⁷ by 23 m.sec. It is tempting to suggest that these constants may have utility beyond the description of this particular set of data but that, of course, is a matter for further experimentation.

Discussion. The tachistoscopic technique as used here appears to be a very stable and sensitive one in which a clear relationship has been demonstrated between prior frequency or familiarity of words and their cognitive time. This further stresses the fact that studies on perception cannot overlook frequency as a major factor in determining visual thresholds. The relationship also may be used to good advantage by investigators of word-frequencies and the related problems

TABLE IV

MEAN RECOGNITIVE THRESHOLD (IN M. SEC.) FOR EVERY INFORMATION-CLASS

Information per symbol (in bits)	2.000	3.322	4.322	5.644	6.644
Recognitive threshold	168	203	221	271	289

of meaning, interests, and values. The technique may offer a means of investigating individual's states of familiarity with verbal symbols which have previously been difficult to ascertain quantitatively.

Whether the rephrasing of the results in terms of information-theory is valuable depends, of course, on whether it leads to new and fruitful hypotheses. It has been suggested by Takada, for example, that research in discriminative reaction-time (which can be predicted by similar equations involving the number of alternatives available) be extended to examine the hypothesis that channel capacity (bits of information per sec.) has a constant value for each individual.⁷ The same suggestion should be examined in this situation. Inspection of Solomon's data and those of this experiment suggests that this hypothesis is not correct, but the necessary experiments should be conducted before this conclusion is accepted.

⁷ Y. Takada, *Critical review of psychological research related to the mathematical theory of communication and a preliminary investigation of the applicability of the theory to psychological data*, Memorandum Report D-4, Task D, Contract No. AF 33(038)-25726, University of Illinois, 1952.

SYSTEMIC FATIGUE AND PERCEPTUAL ORGANIZATION

By WILLIAM A. LYBRAND, T. G. ANDREWS, and SHERMAN ROSS,
University of Maryland

The concept of fatigue has proved to be controversial and troublesome.¹ Investigators have been urged to return to observations by and of Ss in a fatigued state for hypotheses which may prove fruitful for further study.² A review of some studies, which provide such observations, revealed a behavioral category which has received relatively little systematic investigation; namely, perceptual organization.³ A consistent feature of the observations reported in these studies seems to be that the fatigued Ss displayed a lack of ability to form relationships among various stimuli or to relate these stimuli to past experiences. A further impetus for exploring some aspects of perceptual disorganization under fatigue-conditions was received from a study by Ehrenstein.⁴ A significant increase was reported in the effect of each of five illusions as measured on shipyard workers before and after an average day's work. The Müller-Lyer Illusion provided the largest increase. Other studies of this type have employed figures of reversible perspective.⁵ The Ss verbal report on the number of reversals which occur for a unit length of time was used as the basic measure. The results obtained suggested that perceptual disorganization may have occurred in these studies also, since increases in the number of reversals were reported after fatigue-inducing conditions.

On the basis of these considerations, we designed an experiment in which various perceptual tasks were related to fatigue in a systematic manner.

Procedure. Initially, seven dependent variables were selected for study. Thurstone's factorial study of perception offered a basis for the original selection.⁶ The tasks

* Accepted for publication December 21, 1953. This study was performed under Contract DA-49-007-MD-222 between the Medical Research and Development Division, Office of the Surgeon General, U. S. Army and the University of Maryland.

¹ Madison Bentley, Tools and terms in recent researches: Is fatigue tiring?, this JOURNAL, 57, 1944, 264 f.; J. L. Finan, J. C. Finan, and L. D. Hartson, A review of representative tests used for the quantitative measurements of behavior-decrement under conditions related to aircraft flight, *USAF Tech. Rept.* No. 5830, Air Materiel Command, July, 1949; J. H. Bowen, T. A. Hussman, and W. A. Lybrand, A review of the literature on induced systemic fatigue, *Tech. Rept.* No. 7, Proj. DA-49-007-MD-222, Army Med. Res. & Devel. Div., from University of Maryland, March, 1952.

² S. H. Bartley and Eloise Chute, *Fatigue and Impairment in Man*, 1947, 1-429.

³ Leonard Carmichael and W. F. Dearborn, *Reading and Visual Fatigue*, 1947, 1-483; A. S. Edwards, Effects of loss of one hundred hours of sleep, this JOURNAL, 54, 1941, 80-91; G. L. Freeman, Suggestions for a standardized stress test, *J. Gen. Psychol.*, 32, 1945, 3-11; Nathaniel Kleitman, F. J. Mullin, N. R. Cooperman, and S. Titelbaum, *Sleep Characteristics*, 1937, 1-87; D. B. Tyler, The fatigue of prolonged wakefulness, *Fed. Proc.*, 6, 1947, 218.

⁴ Walter Ehrenstein, Die optischen Täuschungen, *Indus. Psychotech.*, 19, 1943-44.

⁵ I. E. Ash, Fatigue and its effect upon control, *Arch. Psychol.*, 4, 1941 (No. 31), 1-81; H. L. Hollingworth, Perceptual fluctuations as at fatigue index, *J. Exper. Psychol.*, 24, 1939, 511-529; Lyle Tussing, Perceptual fluctuations of illusions as a possible physical fatigue index, *ibid.*, 29, 1941, 85-88.

⁶ L. L. Thurstone, *A Factorial Study of Perception*, 1944, 1-148.

chosen for the measurability of the organizational elements involved were: (1) Perception of Hidden Figures; (2) Müller-Lyer Illusion; (3) Kohs Block Designs; (4) Primary Mental Abilities (Space Battery); (5) Size Constancy; (6) Shape Constancy; and (7) Perception of the Upright.

Generally speaking, perceptual organization may be measured in a direct or an indirect manner. In the direct type, the organization *S* ordinarily maintains is measured, as in constancy studies. In the indirect type, *S* is required to adjust his organization to some requirements or distractions of the task-situation, such as is involved in the perception of hidden figures. A pilot study, in which the performance of one group of *Ss* was measured early in the morning and of another group late in the evening, suggested that tasks involving readjustment or organization would be more sensitive to variations of independent variables similar to the one used in the pilot study.

The two independent variables, sleep deprivation and physical exercise, used in this investigation were selected on the basis of their applicability to a wide variety of situations. An initial criterion of selection limited the choice to conditions which are ordinarily assumed to induce general systemic fatigue, rather than specific sensory or motor fatigue. Thus, it was expected that behavioral variations noted under these conditions would be of interest to those individuals concerned with the theoretical aspects of fatigue and also of practical interest to those who are involved in operations similar to the environmental situations employed.

A basic 2×2 factorial design was used for each of the tasks. An absence and presence condition for each variable was employed as indicated in Table I. The factorial arrangement made it possible to obtain an estimate not only of the effects of each of the independent variables, but their combined effects as well.

Performance on the Kohs Block Designs, the Perception of Hidden Figures, and the Müller-Lyer Illusion was analyzed for four groups of *Ss*. Each group was placed under one of the four possible combinations of independent variables as identified by the cells of the 2×2 design. The physical exercise (Exercise in Table I) was a five-mile march with a 40 pound pack. The marching was done at a rate of 120 steps per minute. Sleep deprivation was the loss of one regular night's sleep (No sleep in Table I). A total of 48 men, college undergraduates, served as *Ss*; being randomly assigned to each cell of the design. The same *Ss* served for all of the tasks. The order of taking the tasks was counterbalanced within each group of *Ss*.

It was anticipated that individual differences would be large. Because within-group differences were to be used statistically to evaluate between-group differences, some control of this source of variability was needed. Accordingly, the scores on the Kohs Block Designs were treated by the analysis of covariance, with pre-experimental measurements on the Cards Section of the PMA Space Battery used to adjust the experimental measurements. The same technique was used with the Perception of Hidden Figures, with experimental scores adjusted by pre-experimental measurements on a different set of hidden figures. The technique was not used for the Müller-Lyer Illusion, since Ehrenstein had demonstrated that when account of individual differences is taken, the size of the illusion increases after physical exercise.

The measure taken on each task for each *S* was: (1) Müller-Lyer Illusion—mean length in inches of variable segment of 6 adjustments to 12 in. standard segment; (2) Kohs Block Designs—mean scores in seconds required to complete

four designs; and (3) Perception of Hidden Figures—mean score in seconds required to locate and trace three simple figures imbedded in more complex colored figures.

Results and discussion. The mean performance score for each of the groups on the tasks is shown in Table I. In the case of the Kohs Block Designs and the Perception of Hidden Figures, the scores are the means after adjustments for individual differences in basic abilities. The effects of the statistical adjustment varied. Generally, the means were increased, with the exception of the group sub-

TABLE I
MEAN SCORES ON ALL TASKS FOR ALL GROUPS

	Sleep		No sleep	
No exercise	M-L Illusion	13.9	M-L Illusion	14.1
	Kohs Block	74.4	Kohs Block	67.6
	Hidden Figures	34.6	Hidden Figures	46.9
Exercise	M-L Illusion	13.6	M-L Illusion	13.9
	Kohs Block	53.4	Kohs Block	64.4
	Hidden Figures	23.3	Hidden Figures	38.8

jected to marching on the Kohs task, and the control group on the Hidden Figures task. There was an average increase in precision of the analyses of 28% due to using the regression technique.

As might be expected from an inspection of the mean values shown in Table I, an analysis of variance of scores on the Müller-Lyer Illusion revealed no significant differences between groups under the various conditions of measurement. Table II presents a summary of the results of the analysis of covariance of scores for both the Kohs Block Design and the Perception of Hidden Figures.

The results, while not conclusive, offer hypotheses for further research. Performance on the Kohs Block Designs (time-scores) was more efficient after physi-

TABLE II
ANALYSIS OF COVARIANCE OF SCORES ON KOHS BLOCK DESIGNS
AND THE HIDDEN FIGURES

Source	df	Kohs Block Designs	Hidden Figures
Between physical exercises	1	4206.7†	1103.7
Between sleep deprivation	1	3001.3*	2220.8*
Interaction: phys. ex. and sleep dep.	1	3190.3*	43.7
Error within groups	43	440.0	386.2

* Statistical significance at 5% level.

† Significance at 1% level.

cal exercise, but less efficient after sleep deprivation. Performance on the Hidden Figures task was also less efficient after sleep deprivation, but was not affected by physical exercise. Performance on Kohs Block Designs was affected differentially by different combinations of the two independent variables.

As a tentative conclusion, we may say that perceptual organization is more efficient after mild physical activity and less efficient after sleep deprivation. Both of these conditions are, however, usually assumed to induce general systemic fatigue. A number of factors might account for these findings. The physical activity used

in the experiment may not have been severe enough nor imposed long enough to have a deleterious effect. The work may have served as a warming-up condition. The increase in efficiency may then be a function of increases in perceptual vigilance produced by this warming-up period.⁷ A severer or longer imposed condition of physical exercise might be expected to induce a decrement in performance. Sleep deprivation would not be accompanied by the postulated increase in perceptual vigilance. Thus performance on tasks requiring a high degree of alertness to a variety of stimuli would be less efficient. Cues from stimuli in the environment would be ignored or related to other stimuli more slowly.

It is interesting to speculate on the discrepancy between the effects of physical exercise on the Kohs Block Designs and those effects on the Hidden Figures task. From our knowledge of the factors involved in the tasks, it appears that the Kohs' task requires more of an induction or problem-solving ability, in addition to more manipulative skill, than is required for the Perception of Hidden Figures.

The results strongly suggest that perceptual organization is a behavioral category which is sensitive to variations in conditions assumed to induce general systemic fatigue. The changes also seem to be in accord with variations in physiological conditions which occur under these conditions. Additional research is called for to ascertain the relationship of perceptual organization to fatigue conditions with respect to the type of conditions used to induce fatigue and the nature of the tasks used to measure perceptual organization.

⁷Increases in respiration, heart beat, blood circulation, and metabolism in general would increase the rate of removing toxic substances. With continued physical activity, the accumulation of these physiological correlates of fatigue would be at a higher rate, and we could thus expect the 'refurbishing' process to be inadequate.

THE EFFECT OF PAST EXPERIENCE ON THE PERSPECTIVE REVERSAL OF A TRIDIMENSIONAL FIGURE

By PAULINE AUSTIN ADAMS, University of California

This study is concerned with the persistent effects of past experience on the organization of a perceptual process. The problem to be investigated is the effect of past experience upon the perception of a tridimensional figure.

We selected a wire cube as our stimulus-object because it can readily be experienced in two different ways and also because the conditions of reversal provide an opportunity for testing the effects of past experience.¹ A reversal of the cube consists in an inversion of the perspective of the near and far faces with an accompanying loss of size-constancy—i.e. the near face appears to be smaller than the more distant face. Furthermore, the version perceived at the outset of an observation is always the one corresponding to the objective position of the cube in space. Monocular reversal of the perspective is achieved by practiced Ss in approximately 5 sec.

The general hypothesis that past perceptual experience influences perception can be tested by determining whether the monocular perception of the reversed version of the cube influences the subsequent perception of it after an interval of no stimulation when the same or the other eye is used in the observation.

Procedure. There were two experimental conditions. Condition 1 involved changing the eye used for monocular viewing of the cube after a reversal had occurred. In Condition 2, S covered the eye for which the reversal had occurred for a period of time corresponding to the time required for shifting eyes in Condition 1.

The skeletal cube, which was painted dull black, was constructed of rectangular bars $5\frac{1}{2}$ in. long and $\frac{1}{8}$ in. thick. Four Ss, experienced in the reversal of the figure, participated in the experiment. All of them were right-handed and right-eye dominant. S viewed the cube monocularly from a distance of 28 in. by focussing at a uniform white background 35 in. behind the cube until a reversal had occurred and persisted for 2 sec. Both eyes were then covered for a 2-sec. period. In Condition 1, S uncovered the eye which had been covered when the reversal occurred. In Condition 2, the same eye for which the reversal occurred was reexposed. S indicated if the cube appeared in its reversed form or in its objective position. These results will be referred to as reverse cube and obverse cube.

Four conditions of observation were investigated: R-L (shifting from the right to the left eye), L-R (shifting from the left to the right eye); and R-R and L-L (in which the same eye was reexposed). Conditions R-L and R-R, in which S began a trial using the right eye will be referred to as right-eye training; similarly

* Accepted for publication August 26, 1953. The author's thank are due Professor Leo Postman for assistance in this study.

¹ The effects of training on the perception of ambiguous two-dimensional stimuli has been demonstrated by M. R. Harrower, Changes in figure-ground perception in patients with cortical lensions, *Brit. J. Psychol.*, 30, 1939, 47-51; and R. Leeper, A study of a neglected portion of the field of learning—the development of sensory organization, *J. Genet. Psychol.*, 46, 435, 4-175.

L-R and L-L will be referred to as left-eye training conditions. The order of these conditions was systematically varied that each condition was represented four times in each of the four series of observations. Observations within a series were separated by at least 15 sec., while the time elapsing between series of observations was at least 2 hr. The parallax which resulted when the viewing eye was shifted could be compensated for by a slight shift of the body. A reduction-screen was found un-

TABLE I
FREQUENCIES OF ALTERNATIVE PERCEPTIONS OF THE CUBE
UNDER VARIOUS EXPERIMENTAL CONDITIONS

Ss	Condition							
	Different eyes				Same eye			
	R-L		L-R		R-R		L-L	
	reverse	obverse	reverse	obverse	reverse	obverse	reverse	obverse
LR	13	3	7	9	12	4	14	2
LP	7	9	7	9	10	6	8	8
DJ	10	6	5	11	13	3	11	5
PA	15	1	11	5	14	2	12	4
Totals	45	19	30	34	49	15	45	19

satisfactory as it permitted partial perception of the cube during the ocular shift. All the Ss served under all the conditions.

Results. The results for each S in each of the four conditions are presented in Table I, which shows the frequency with which the Ss perceived the cube in obverse and reverse form following the initial reversal. To test the significance of the difference in the number of reversals under the different experimental conditions, an analysis of variance was performed (see Table II).

The *F*-ratios show that the differences between Conditions, Types of Training, and Ss are significant at better than the 5-% level. The interaction between conditions and training fails to reach significance. All conditions of preliminary training

TABLE II
ANALYSIS OF VARIANCE OF REVERSAL SCORES

Source	df	Mean square	F
Conditions	1	22.56	7.50*
Training	1	22.56	7.50*
CXT	1	7.57	2.51
Ss	3	18.73	6.22*
Error	9	3.01	—

* $F = 5.12$ at the 5-% level, 10.56 at the 1-% level for 1 and 9 degrees of freedom.

lead to a highly significant number of reversals when compared with 100% initial perception of the cube in the unreversed position.

Discussion and summary. The general hypothesis of an influence of past experience on the perception of a tridimensional figure is verified by the demonstration that once a reversal of a cube has occurred, the perceptual process corresponding to the reversed version of the cube can persist in the absence of further retinal stimulation

and can influence the subsequent perception of the figure. This influence of past experience on the perception of a wire cube is demonstrated by the significant number of perceptions of the reversed version of the cube after the training procedures, as compared with 100% initial perceptions of the cube in its objective position. Our interpretation is supported by the Ss' observation that when they initially perceived the cube in its objective position after the interval of no stimulation, the reversal time was very rapid.

The significance of the difference between R-L and L-R as compared with conditions R-R and L-L is interpreted as evidence for the fact that the locus of the prior stimulation is indeed a significant variable. (We assume that this result is not an artifact of the slight parallax involved in the ocular shift.) The significant difference between the right and left training conditions may be a consequence of cortical dominance. The process which has its locus predominantly in the left hemisphere might exert a greater influence on the subsequent perception.

CRITICAL FLICKER FREQUENCY, AGE, AND INTELLIGENCE

By CARROLL M. COLGAN, University of Florida

Studies by Misiak and others have shown that elderly people have by and large lower critical flicker frequencies (CFF) than younger ones.¹ Studies by Hartmann² and Miller,³ on the other hand, found no significant difference in CFF between children and young adults. These results considered together suggest that the CFF may measure the process of aging. If the hypothesis is true, then the higher age-ranges should exhibit a marked decline in CFF. This paper presents measurements of CFF of two groups of elderly Ss.

Subjects. All the Ss were men residents of Moosehaven, the home for aged members of the Moose Lodge, in Orange Park, Florida. There were 18 Ss in Group I (65-80 yr. old) and 22 in Group II (81-95 yr. old).

Apparatus. The light source was a glow modulator tube, Sylvania R-1130, modulated by an electronic switch. The wave-form was essentially rectangular, thus limiting the latency and decay of the light source to the time required for ionization and de-ionization of the gas contained therein which was approximately 5 μ sec.

The stimulus-patch was white in color, circular in shape, 6.5 mm. in diameter, and 3 ft. from S's eyes. The intensity of the patch was 63.50 foot-lambert and, when flashing, had a light-dark ratio of 1:1. The light and test-patch were enclosed in a large rectangular box, the insides of which were flat-black.

Procedure. The procedure was the same for all Ss. After adapting 10 min. in a partially darkened room, S fixated the patch binocularly. Flicker and the fusion-point were given ostensive definition for the S. Five ascending and five descending trials were given S who was instructed to report when the flicker disappeared and reappeared. The rate of acceleration was high and irregular. The initial rates of repetition were randomized and the exposure was continuous for each trial (approximately 6 sec.).

Results. The mean CFF for each group was determined and a test of the signifi-

* Accepted for publication October 5, 1953. This research was done while the author was a Public Health Service Research Fellow of the National Institute of Mental Health. The author wishes to express his thanks for the coöperation of Dr. R. W. Kleemeier, the Director of the Moosehaven Research Laboratory for Gerontology.

¹ Henryk Misiak, Age and sex differences in critical flicker frequency, *J. Exper. Psychol.*, 37, 1947, 318-332; Josef Brožek and Ancel Keys, Changes in flicker fusion frequency with age, *J. Consult. Psychol.*, 9, 1945, 87-90; P. W. Miles, Flicker fusion fields: Effect of age and pupil size, *Amer. J. Ophth.*, 33, 1950, 769-772; Ernst Simonson and Josef Brožek, Flicker fusion frequency: Background and applications, *Physiol. Rev.*, 32, 1952, 349-378.

² G. W. Hartmann, Comparison of the flicker thresholds in children and adults, *Child Develop.*, 5, 1934, 122-126.

³ V. L. Miller, Critical frequency limit for visual flicker in children between the ages of six and eighteen years, *Genet. Psychol. Monog.*, 26, 1942, 3-55

cance of the difference between the two computed. This difference was found to be significant at the 10-% level of confidence ($t = 1.7$, $p = 0.10$). There is accordingly one chance in ten that the difference in *CFF* actually observed between the two age-groups could have been due to sampling errors. The difference between the two groups is of slight significance. This result is surprising because the differences obtained in the earlier studies were quite significant.⁴

The work of Tanner suggests that perhaps the two groups differed in more than age.⁵ In a study with college students, he found a significant relationship between *CFF* and the ACE Psychological Examination. Since this test presumably measures intelligence as well as college aptitude, there is consequently the possibility that this factor was responsible for the observed difference in *CFF* in the present study. To test this possibility we obtained the *Ss'* scores on the Wechsler-Bellevue Intelligence Scale.⁶ The means for the two groups were computed (using total scores) and the

TABLE I
CFF vs. VARIOUS WECHSLER SCORES PRODUCT-MOMENT CORRELATION COEFFICIENTS

Verbal		Performance	
(N=40)			
Subtest	r	Subtest	r
Information	+0.20	Picture Arrangement	+0.35*
Comprehension	+0.30*	Picture Completion	+0.44†
Digit span	+0.03	Block Design	+0.45†
Arithmetic	+0.14	Object Assembly	+0.36*
Similarities	+0.44†	Digit Symbol	+0.39*
Vocabulary	+0.31*		
Verbal Total (prorated)	+0.32*	Performance Total	+0.48†
Wechsler-Bellevue Total Score			+0.46†

* Significant at the 5-% level of confidence.

† Significant at the 1-% level of confidence.

test of significance of the difference carried out. This difference was found to be significant at the 1-% level of confidence ($t = 3.27$, $p = 0.01$).

It now became important to determine, whether age or intelligence was the pertinent variable in the comparison of the *CFF*. To answer this question, correlation coefficients were computed for the following: (a) *CFF* vs. age; (b) Wechsler-Bellevue (*W-B*) vs. age; (c) *CFF* vs. *W-B*; (d) *CFF* vs. age with the *W-B* held constant; and (e) *CFF* vs. *W-B* with age held constant. These results are given in Table 2.

Inspection of this Table shows that the hypothesized relationship of *CFF* with intelligence is sustained. It is clear that part of what is measured by the Wechsler-Bellevue Scale is also measured by *CFF*—at least for our *Ss*.

⁴ Brožek and Keys, *op. cit.*; and Miles, *op. cit.*

⁵ W. P. Tanner, Jr., A preliminary investigation of the relationship between visual fusion of intermittent light and intelligence, *Science*, 112, 1950 (No. 2903), 201-203.

⁶ These scores were made available to the author by Mr. Robert W. Arms who was currently engaged in a longitudinal study. All of the *Ss* of the present study were drawn from his much larger group.

The comparison of *CFF* with age gave a product-moment r of -0.32 , significant at the 5-% level. The *W-B* was also significantly related to age, $r = -0.53$ with $p = 0.01$. The partial correlation of *CFF* and age with *W-B* held constant gave an r of -0.10 ($p = 0.42$). On the other hand, the partial correlation of *CFF* and the *W-B* with age held constant produced an r of $+0.36$ ($p = 0.05$).

Conclusions. The following conclusions may be drawn from these results: (1) the difference in *CFF* between Groups I and II of this study was probably due to intelligence, not age; (2) the results of this study, in corroboration of those of Tanner's demonstrate that there is a significant relationship between *CFF* and intelligence which is independent of age; and (3) previous studies relating *CFF* to age must be reexamined to determine whether intelligence was a controlled variable—if it was not, the conclusions of those studies must be qualified accordingly. It may be that the relation of age to intelligence, and to *CFF* when intelligence is not a controlled factor, stems from the fact that after the peak of development is reached, the passage of time provides opportunity for the accumulation of physiological accidents or events that contribute bit by bit to the impairment of neural functions upon which both intelligence and *CFF* depend.

The subtests that correlated lowest with *CFF* could possibly have done so by being unreliable tests for these *Ss* and this could have resulted from the *Ss'* lack of attentiveness. Those tests that give *S* some material object he can attend to (the performance subtests) gave significant coefficients; this does not, however, account for the high value for the similarities subtest which is entirely verbal. We have no explanation for the correlational pattern obtained, but we believe that the speculation on attentiveness may prove fruitful in further research on this problem.

PREFERENCE IN THE RAT FOR VERTICAL OR HORIZONTAL STRIPES AFTER TRAINING ON A WHITE-BLACK DISCRIMINATION

By O. THOMAS LAW, The Johns Hopkins University

The present experiment grew from efforts by the writer and others in this laboratory to develop complex processes in the rat. On a long series of visual discriminations such as were used in these attempts, these animals persistently tended to utilize and transfer any available brightness cues. In attempting to control and eliminate such cues we observed certain tendencies of the Ss to respond to stimulus-objects of presumably equal brightness as though they were in fact unequal. The familiar horizontal vs. vertical striations problem seemed particularly susceptible. Accordingly the following experiment was devised and executed to determine whether this problem offered unsuspected brightness cues.

Procedure. Fourteen 90-day old female rats of a hooded strain were trained in an octagonally shaped jumping stand, so designed that eight equally-illuminated windows, each 6-in. square, were placed 8 in. from a centrally placed jumping platform. The Ss were randomly assigned to two groups of seven each. One group, Group W, was trained to jump through the window covered by a white card and to avoid the seven windows covered by black cards; the other group, Group B, was trained on the reverse problem of jumping through the window covered by a black card and of avoiding the seven windows covered by white cards. All the Ss were given 10 trials per day. Food reward (wet mash) was placed behind the correct window and each S was allowed a few bites for each correct jump.

After training to a criterion of 10 consecutive correct responses, each S was rested one day, trained again to a criterion of five correct, rested another day, and finally trained again to five correct responses. Stability of the response was apparent from the fact that no further errors were made by an S after attaining the first criterion of 10 correct trials. Following this pretraining procedure, the Ss were then rested 1 day and tested for 2 consecutive days, 10 trials per day, on striped patterns. These patterns consisted of alternate black and white stripes of 20 mm. width and were alternately arranged in the 8 windows of the jumping stand, 4 vertical and 4 horizontal. On the first test-day, the vertical patterns were placed in the odd-numbered windows; on the second test-day, they occupied the even-numbered windows. All eight windows were unlocked on both days and S was rewarded for each jump regardless of the window chosen.

Results. Results for each of the two test-days are presented in Table I. Responses to vertical stripes only are recorded for Group W, and similarly, responses to horizontal stripes only for the Ss trained on black-positive, since preferences in

* Accepted for publication October 9, 1953. This study was conducted during the tenure of a Post-Doctoral Fellowship of the National Science Foundation.

general^o were in these directions. Since there are always 10 trials in each test, responses to non-preferred stimulus-patterns may be obtained by subtracting each score from 10.

Of those in Group W, 5 out of 7 Ss preferred one stimulus pattern to the other by a ratio of 7:3 or more on the first test-day. All 7 Ss in Group B preferred one pattern to the other by 8:2 or more on their first test-day. Defining preference by these ratios, we can say that 3 Ss of Group W preferred vertical stripes and 2 of them horizontal stripes; 6 Ss of Group B preferred horizontal stripes and 1 preferred the vertical. The second test-day showed some slight shifts in these preferences: One S from Group W changed from vertical to horizontal preference, one from Group B did the reverse, and another from Group B shifted from a horizontal preference to none at all.

^o Entering the data of Table I in a 4-fold table, the following values of χ^2 result: Total $\chi^2 = 29.50$; χ^2 for the first test-day = 27.13; χ^2 for the second test-day = 2.37. The *P* values for total and for the first test-day both lie beyond the 0.1-% level; the *P* value for the second test-day is 0.13 approximately.

As a double-check, the non-parametric Wilcoxon *t*-test was applied to the data of Table I.¹ The *t*-value for the second day did not reach significance. For the total, it

TABLE I
RESPONSES OF Ss PRETRAINED IN WHITE POSITIVE (GROUP W) AND BLACK POSITIVE (GROUP B) TO VERTICAL AND HORIZONTAL STRIATED STIMULI ON 2 SUCCEEDING TEST DAYS

Ss Group W	Responses to vertical		Ss Group B	Responses to horizontal	
	1st day	2nd day		1st day	2nd day
1	10	10	2	10	10
4	1	1	3	10	10
5	3	3	6	10	10
7	7	1	9	9	4
8	5	4	10	8	7
11	10	10	12	9	3
14	6	4	13	2	1
Total	42	33		58	44

fell just short of the 5-% level. The data for the first test-day yielded a *t*-value significant between the 5-% and 2-% levels. These values indicate a significant preference of Group B for horizontal stripes, and of Group W for vertical. Most of the significance of these data is contributed by Group B since Group W's responses, while almost equally marked, were less uniform in direction. An examination^o of the initial response in the test-situation for each S yields similar results. Five of the Ss of Group W chose vertical and two chose horizontal patterns on the first jump; five of those of Group B responded initially to the horizontal card and two to the vertical. From the binomial, the combined probability of initial responses by each of the group to five or more of the opposite stimulus-patterns may be found to be 0.05.

The decline in preferences which appeared on the second test-day is probably in part due to the association of reward with response to *any* pattern, and in part to

¹ Frank Wilcoxon, Some rapid approximate statistical procedures, American Cyanamid Co., 1949.

development of some slight window preferences. These latter, however, did not become sufficiently marked in any case to account for 50% of the total responses for the 2 days.

Discussion and conclusions. The conclusion appears justified that Ss have preferences between vertical and horizontal stripes after they have been trained on a black-white discrimination, and that in some cases these preferences are quite pronounced and persistent. In general, those trained to white positive prefer vertical stripes and those trained to black positive prefer horizontal. The reasons for this are not apparent. It is possible that the rat is subject to some sort of vertical-horizontal illusion in viewing these patterns. However the results may be interpreted they make it apparent that some transfer to striped patterns is possible as a result of prior training on problems involving flux or brightness cues.

Summary. Data have been presented on 14 female rats, half of whom were pre-trained on white positive vs. black negative stimulus-objects in an octagonal jumping stand, the other half being trained on black positive vs. white negative. All were then exposed for 20 trials to alternately vertical and horizontal black and white striations in a 100% reward situation. It was noted that those trained on white positive tended to prefer the vertical striations, while those trained on black positive tended to prefer the horizontal striations. The association was significant at a high level of confidence. Better than 80% of the animals tested were noted to show some preference irrespective of direction.

A FURTHER STUDY OF DIFFERENTIAL AFFERENT CONSEQUENCES IN NONDIFFERENTIAL REINFORCEMENT

By F. T. CRAWFORD, G. L. MAYES, and M. E. BITTERMAN,
University of Texas

In the latest of a series of studies of perceptual development during nondifferential reinforcement, a situation was employed in which the difference between the two members of a pair of nondifferentially reinforced stimulus-cards was linked consistently with a striking difference in the afferent consequences of response (goal-box color) which had in itself no motivational significance.¹ The hypothesis was tested that this linkage would maintain an orientation to the difference between the stimulus-cards and thus facilitate the subsequent discrimination between them. The results of the experiment were negative, but there was evidence that the critical discrimination chosen for the work (wide *vs.* narrow vertical stripes) was unsuitable. In the present experiment, the validity of the hypothesis was investigated with more appropriate stimuli.

Subjects. The Ss were 46 experimentally naïve white rats, bred in the laboratory and 3-4 mo. of age at the beginning of the experiment.

Apparatus. A two-window jumping apparatus was employed. Behind the windows was a feeding platform divided by partitions into three goal-boxes, the first and third of which were painted black and the second white. Since the position of the platform relative to the windows could be shifted, it was possible to have either a white box behind the left window and a black box behind the right window or a black box behind the left window and a white box behind the right window. Below the windows was a cloth net into which the animals fell after responses to locked cards. With the exception of the goal-boxes, the entire apparatus was painted mid-gray.

Two black-and-white striped stimulus-cards were employed. One of the cards was horizontally and the other vertically striped. The width of each stripe on each card was 0.5 in.

Preliminary training. Throughout the experiment the animals were maintained on a 24-hr. feeding schedule. After adjustment to handling they were fed wet mash for equal periods of time in white and black goal-boxes. Then they were trained to jump gradually increasing distances (to a maximum of 9 in.) from the jumping platform through the open windows, each jump being rewarded with a nibble of wet mash. Frequency of response to the two box-colors and to the two windows was equated by shifting the positions of white and black boxes and by the alternation of free with manually guided trials. On the free trials most of the animals responded primarily in

* Accepted for publication August 14, 1953.

¹ B. A. Billingsley, W. E. Feddersen, and M. E. Bitterman, Discrimination following nondifferential reinforcement with differential afferent consequences, *THE JOURNAL*, 67, 1954, 335-337.

terms of position, and on this basis they were divided into two matched groups—Group I with 24 animals and Group II with 22.

Experimental design. In the first stage of the experiment, which lasted 12 days, all the animals were given eight nondifferentially reinforced trials per day. On each trial one window held a vertically striped and the other a horizontally striped card (the lateral arrangement of the two cards being systematically changed from trial to trial), and the use of both free and guided trials provided equal experience with each card in each position. Every response was rewarded with a nibble of wet mash. For the animals of Group I, half the responses to each card in each window carried the animal into the white goal-box and the remaining responses led to the black goal-box. For the animals of Group II, there was a consistent relation between card and goal-box color; for half the animals the white goal-box was always behind the horizontal stripes and the black box behind the vertical stripes, while for the remaining animals the reverse relation obtained. The design of the previous experiment on differential afferent consequences involved a third group of animals nondifferentially reinforced on gray cards, which made it possible to determine whether there was any perceptual development at all as a consequence of nondifferentially reinforced experience with the critical stimuli (wide and narrow vertical stripes). In the present

TABLE I
PERFORMANCE OF THE VARIOUS SUBGROUPS

	IA	IIA	Diff.	IB	IIB	Diff.
Initial errors	10.0	11.8	-1.8	11.6	11.0	0.6
Total errors	21.3	24.2	-2.9	22.4	20.0	2.4
Days to criterion	5.4	6.1	-0.7	6.2	5.0	1.2

experiment no such group was necessary since it has already been demonstrated that nondifferentially reinforced experience with horizontal and vertical stripes may facilitate subsequent discrimination between them.²

From the preferences for cards and positions displayed by the animals on free trials during the first stage of the experiment, several subgroups were selected. Groups IA and IIA were made up of 16 matched pairs of animals from Group I and Group II which showed consistent positional preferences but no card-preferences. Groups IB and IIB (which contained, respectively, five and six animals) were made up of those animals from the parent groups which showed card preferences (in most cases for the horizontal stripes). During the second stage of the experiment all groups were trained on the horizontal-vertical discrimination. Half the animals in Groups IA and IIA were trained to jump to vertical and the remaining animals were trained to jump to horizontal. All animals in Groups IB and IIB were trained against the card-preferences previously displayed. During this training, one of the goal-boxes was consistently paired with the positive card. For the animals of Groups IA and IB the goal-box was that which had been consistently paired with the now positive card during the first stage of the experiment; for each animal in Group IIA the box was that used for the paired animal of Group IA; for the animals of Groups IB and

² Bitterman, A. D. Calvin, and C. B. Elam, Perceptual differentiation in the course of nondifferential reinforcement, *J. Compar. & Physiol. Psychol.*, 46, 1953, 323-397; Bitterman and Elam, Discrimination following varying amounts of nondifferential reinforcement, this JOURNAL, 67, 1954, 133-137.

IIB the box was chosen randomly. All animals were given eight trials per day by the method of correction. Each trial consisted of a series of one or more responses to a given pair of cards which terminated with a correct response; a maximum of three errors was permitted on each trial, following which the animal was guided in the correct direction. The criterion of learning was two successive errorless days.

Results. The principal results of the experiment are summarized in Table I. Differences between Groups IA and IIA (tested by Wilcoxon's nonparametric method for paired replicates³) proved to be statistically unreliable, as did the differences between Groups IB and IIB (tested by Festinger's nonparametric method⁴). The consistent linkage of the difference in goal-box color with the horizontal-vertical difference apparently did not facilitate discrimination in the animals of Group I.

Conclusion. The hypothesis which this experiment was designed to test may be rejected.

³ Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures*, American Cyanamid Co., Stamford, Conn., 1949, 1-16.

⁴ Leon Festinger, The significance of a difference between means without reference to the frequency distribution function, *Psychometrika*, 11, 1946, 97-105.

REVERSED DISCRIMINATION AS A FUNCTION OF THE NUMBER OF REINFORCEMENTS DURING PRE-TRAINING

By J. M. WARREN, University of Oregon

Given a pair of stimulus-objects, *A* and *B*, continuity theory predicts that the number of errors made in learning to select *B* will be proportional to the number of times *A* has been reinforced in antecedent pre-training.¹ Results in conformity with this hypothesis have been obtained in several experiments with rats and chimpanzees.²

Although Harlow found that monkeys can learn to solve reversed discriminations in a single trial after 7 to 11 pre-reversal trials,³ there are no data regarding the relation between the number of reinforcements given during pre-training and the number of errors made after reversal. The purpose of the present investigation was to determine the number of errors made by monkeys in learning *B* (positive) versus *A* (negative) after varying amounts of reinforcement on *A* (positive) versus *B* (negative).

Subjects. Three young, experimentally naïve Rhesus monkeys served as *Ss* in this study.

Apparatus. The Wisconsin General Test-Apparatus and the specific testing procedures associated with the use of this apparatus, which are described elsewhere,⁴ were used in this study. The stimulus-objects (108 pairs), varying in multiple visual dimensions, were similar to those customarily used at the Wisconsin Primate Laboratory.

Procedure. Using the non-corrective technique, the *Ss* were tested until they attained a criterion of 1, 2, 4, 8, 16 or 32 correct responses to the *A*-member of a pair of stimuli. Immediately after the criterial response, the *B*-member of the pair became the rewarded object, and testing was continued for another 10 trials. Eighteen reversed discriminations were tested following each of the six degrees of reinforcement during pre-training; thus, a total of 108 problems was presented to each animal.

Three problems were presented each day for 36 days, and amount of reinforcement during pre-training was counterbalanced by use of latin squares.

* Accepted for publication September 24, 1953. This research was conducted at the USAF School of Aviation Medicine Primate Laboratory, Austin, Texas.

¹ K. W. Spence, Analysis of the formation of visual discrimination habits in the chimpanzee, *J. Comp. Psychol.*, 23, 1937, 77-100.

² Spence, *op. cit.*, 87; T. L. McCulloch and J. G. Pratt, A study of the pre-resolution period in weight discrimination by white rats, *J. Comp. Psychol.*, 18, 1934, 271-290; D. Ehrenfreund, An experimental test of the continuity theory of discrimination learning with pattern vision, *J. Comp. & Physiol. Psychol.*, 41, 1948, 408-422; F. P. Gatling, A study of the continuity of the learning process as measured by habit reversal in the rat, *ibid.*, 44, 1951, 78-83.

³ H. F. Harlow, Performance of Catarrhine monkeys on a series of discrimination reversal problems, *ibid.*, 43, 1950, 231-239.

⁴ Harlow, The formation of learning sets, *Psychol. Rev.*, 56, 1949, 51-65.

Results. The results are presented in Fig. 1, which shows the percentage of correct responses after reversal as a function of the number of reinforcements during pre-training, with blocks of 36 problems as a parameter. It is apparent that performance improved markedly between successive thirds of the experiment. The changes in performance are consistent and highly significant, although the F for differences among Ss is not significant. The F for the significance of the differences

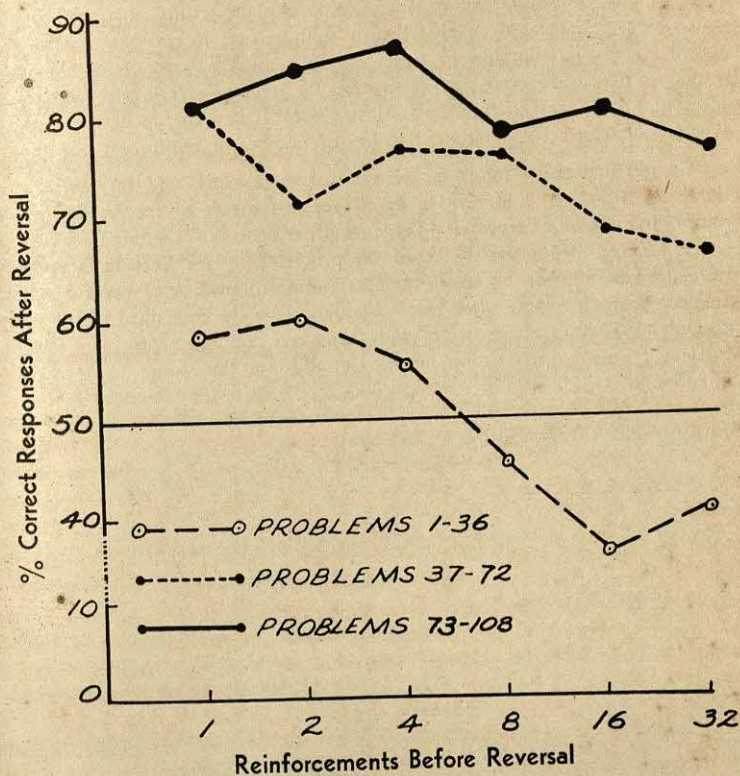


FIG. 1. PERFORMANCE AFTER REVERSAL AS FUNCTION OF THE NUMBER OF REINFORCEMENTS BEFORE REVERSAL

among blocks of problems (thirds) is 120.39 ($P < .01$ with 2/20 df). These data indicate a large amount of inter-problem transfer—the formation of a learning set.

The changes in the shape of the functions are more significant theoretically than the overall increase in efficiency of performance. On problems 1 to 36, the difference in the percentage of correct responses following 1 and 32 reinforcements during pre-training was 21%. The difference was only 5% on problems 73 to 108. Analyses of variance indicate that the variance due to the number of reinforcements during

pre-training was significant only on problems 1 to 36. This variable had no significant effect on performance during the last two-thirds of the experiment.

Thus, the results of this study corroborate those obtained by Harlow, and fail to support the implications of either continuity or non-continuity theory. Spence's statements regarding the effect of reinforcement during pre-training on reversed discriminative learning, which are presumably meant to apply to *Ss* at all stages of training, are obviously not confirmed by the results of this experiment. As Harlow pointed out, non-continuity theory is unable to explain results such as those here presented since it has been developed only with respect to reversals before *S* has developed systematic responses to either stimulus-object.⁶ Furthermore, the virtual independence of post-reversal performance on the number of reinforcements during pre-training, which was attained in the last phase of the experiment, was preceded by a period of significant determination by the number of reinforcements.

Summary. Three naïve monkeys were tested on a series of 108 reversed discriminations. Either 1, 2, 4, 8, 16, or 32 reinforcements of the finally negative stimulus-object preceded reversal. Large improvements in the accuracy of performance occurred from one block of 36 problems to the next throughout the experiment. The number of reinforcements during pre-training was a significant determinant of performance after reversal only during the first third of the experiment. Only 5% more errors were made after 32 reinforcements than after one reinforcement during pre-training in the last third of the study. The results fail to support either the continuity or non-continuity theories of learning.

⁶ Harlow, *op. cit.*, 238.

NOTES AND DISCUSSIONS

ON VERIFYING HYPOTHESES BY VERIFYING THEIR IMPLICATES

* Let there be given a class of propositions comprizing $H, I_1, I_2, I_3, \dots I_n$ in which the material truth of H is undetermined while each of the I -propositions has been proved to be materially true. In this situation, each of the I -propositions is a material implicate of H . We may assert them all conjointly, so as to yield one compound implicate I_c . We may then truly say that if H is true, I_c is also true; but we can not assert that because I_c is true, H is true, for to do so would be to assert the consequent and by the rules of logic to get no valid conclusion at all. The truth of H depends on the truth of I_c (and therefore on the truth of every component of I_c) in so far as H can not be true when I_c is false; but the material truth of I_c is independent of the truth of H .

This fallacy is readily seen when it is expressed in arbitrary symbols; it may or may not escape detection when it is expressed in ordinary words. For example:

- (1) If 'Old Dog Tray' was run through a large and powerful sausage-grinder, he is dead; he is dead, therefore, he was sausaged.
- (2) If Plato died of typhus fever, he was bitten by at least one insect; he was insect-ridden, therefore, he died of typhus fever.
- (3) If a person is intoxicated, he staggers, speaks thickly, and smells of liquor; this person staggers, speaks thickly, and smells of liquor, therefore, he is intoxicated.
- (4) If a person has an over-compensated inferiority complex, he blusters, is aggressive, domineering, and dogmatic; this man blusters, is aggressive, domineering, and dogmatic, therefore, he has an inferiority complex.

All these specimens of para-reasoning have the same form: they all assert the consequents, and the conclusions are invalid even though some of them may be true to the facts of perceptual observation. When we encounter such an example, do we not often challenge the material truth of the conclusion without examining the reasoning processes at all? I dare say that we all do. Thus, we might reject the first conclusion as being incredible because we doubt the proverbial recipes for wienerwurst and frankfurters. We might accept or reject the fourth according to our liking

or aversion for the person in question. All these grounds, however, are psychological, not logical. We fall into some attitude or set that prevents us from using logic.

It is easy enough to examine the postulates of a two-valued logic, and by reasoning, intuition, or feeling convince ourselves that each of its theorems is valid; but to discern relations among universals is one thing, to discover that a particular instance exemplifies such a relation is quite another. Proficiency in both these abilities is necessary to skillful reasoning—proficiency in either without the other is insufficient. Perhaps it may turn out that different learning procedures or different training procedures are called for in order to achieve both proficiencies.

If the hypothesis *H* is fertile and simplifying, then one may find it hard to doubt its truth, even though only a few of its implicates have been detected, examined, and found to be materially true. One's feeling of assurance does not, however, guarantee the material truth of any proposition. Even though some mystics, romanticists, and religionists assert that it does, all scientists, while working at their trade, must deny that it does.

We need mention as reminders only a few hypotheses that some noted scientists have—so they believe—verified by verifying some of their consequents without falsifying others. Any hypothesis concerning the initial condition and the laws of change in the earth's structure will suffice; so will any fashionable hypothesis concerning biological evolution, stellar evolution, expansion of the universe, 'the' structure of 'the' atom, thinking and bodily movement (as set forth by G. H. Lewes, William James, Knight Dunlap, and John B. Watson). Every sophisticated scientist knows that these hypotheses are fictions; we may entertain them so long as we find no falsifiers, but we need not entertain them; and we may discard them whenever we find others that are more fertile or more simplifying. So went, for example, the phlogiston hypothesis; so went the Ptolemaic hypothesis concerning planetary motion. Neither hypothesis has yet been shown to imply any falsehoods; but more convenient ones were invented, and the older, more cumbrous ones fell into disuse.

All logicians—even mere clever dabblers in logic—agree that inductive reasoning is based on affirming the consequents. Many of them, however, hesitate to rate the procedure as invalid. They say of some hypotheses, as the Freudian says of his hypothesis of the meaning of dreams, that it 'works.' They overlook the point that if an hypothesis is false, exactly half of its implicates will be true. Thus, they manifest certain reaction-patterns that are at least curious.

Some propose to verify an hypothesis by examining all its implicates and verifying each of them by observation. But how many implicates does it have? Let us recall that if an hypothesis H is materially false, it implies all the propositions there are. Call their number N . If H is true, it materially implies all the true propositions there are. Their number is $N/2$. It seems to me to be obvious that N is certainly very large, and that $N/2$ is large also. Indeed, for all that we know, N , the total number of propositions, expressed or unexpressed, may be infinite. If it is infinite, so also is $N/2$. If it is finite but still very large, then the task of detecting, examining, and verifying every implicate of H becomes inconveniently heavy, and it may be, infeasible. If N is infinite, the task becomes impossible.

Some writers have said—I think incautiously—that by discovering any previously undetected implicate and verifying it by observation, we increase the probability that H is true. This is not so; we merely intensify our expectational belief that any implicate we detect and examine hereafter will be true also. The intensity of one's expectation or the strength of his faith is, however, different from the probability that the event which he is contemplating will occur. For example, consider a deal in whist which has just been completed. What is the probability that each of the players has been dealt the hand that he actually got? It is of the order of 10^{-27} . In other words, there are a million billion trillion possible distributions of 52 cards, 13 each to each of 4 players. Any of these many distributions was equally likely to occur; this distribution did occur; the dealer brought off a very improbable event. Now, if in this deal, every player had obtained a perfect hand, the fact would be good for head-line news stories. The latter distribution is no more improbable, however, than the one actually dealt is, yet the one actually dealt does not astonish us. It would astonish us if our expectation depended on the probability of the occurrence of the event. Indeed if the occurrence of improbable events constitutes a miracle, then perhaps every event we participate in is miraculous.

* What are hypotheses for? For what, indeed, except as impliers to enable us to deduce a set of implicates that are materially true, interconnected, and useful. It is only from materially true hypotheses that we can strictly deduce materially true implicates. We have, therefore, to verify H in order to use it as an implier. If we are to verify it by verifying its implicates, why not verify the latter and desist?

Some have proposed that to verify an hypothesis by verifying its implicates, one need not examine and verify all, but only as many as convenient. Following the examples of Roger Bacon, J. S. Mill, and Jevons,

one now executes what some have called an 'inductive leap' and *assumes* that all the undetected, unexamined, and unverified implicates are true. This begs the question, as does every other extrapolation.

A Missouri farmer, called affectionately by his neighbors 'Uncle Heeb,' once gave two lucid examples:

(1) When the first day of March rolled around each year, he pronounced that he was safe until the end of December because he had noticed that whenever he had lived until the first of March he had lived through to the end of the year. The facts agreed with his prediction for 85 years, but failed in his 86th year.

(2) Uncle Heeb became bored with reading about the preparations for the Byrd expedition to the South Pole, especially by the precautions taken against the cold. They were, in his opinion, the silliest lot of men he had ever heard of. He himself had received little formal schooling because he was orphaned during the siege of Vicksburg; but he had tried to make good use of what he had by noticing what went on about him. As a buyer of cattle, he had traveled as far north as central Saskatchewan Province, Canada, and as far south as Old Mexico. The farther north he went, the colder it got; the farther south he went, the hotter it got. Everybody knows that the North pole is the coldest spot on earth; the South Pole must, therefore, be the hottest.

It is easy to question this reasoning, for the special circumstances tend to alert the reader or hearer, and to cause him to suspect, in the second instance at least, that some premise the truth of which may be questioned has been suppressed. In principle, however, the reasoning is exactly like that which is followed in the detection of business-trends, the estimation of the population of a given country half a century hence, etc. My better undergraduate students recognized this analogy, and a few of them renamed extrapolation and the inductive leap "Uncle Heeb-ing."

In brief, then, so called inductive reasoning is essentially affirming the consequents. In practical situations it is not at all rare, and in some instances it has been appallingly expensive. For example: Strong obtained from many successful engineers a set of answers whether they liked, disliked, or were indifferent to certain activities or vocations. These engineers agreed, so it is said, in choosing similar patterns of interests. It was urged *therefore*, that from applicants for admission to engineering colleges, those who were accepted should manifest these interests.

Again, some authors have asserted truly that if n accidents should befall N operators without regard to the personal identity of any of them (and therefore without regard to any of their personal characteristics), then the distribution of accidents per operator will follow Poisson's law. But it does not follow therefrom that if the distribution is Poissonian, then every operator is as liable to accident as every other, although some recently have tendered this conclusion, based on affirming the consequent. Moreover, if a distribution is Poissonian, its mean and the variance therefrom will

be equal. These same investigators presented some distributions, which they did not describe in detail, in which the variance equalled the mean. From this they concluded that the distribution was Poissonian without testing it for closeness of fit; and *therefore*, that in the sample population there were no personal traits affecting one's liability to accident.

Again, in psychoanalysis, when the patient exhibits certain patterns of anxiety and related symptoms, it is usual to say that if, in his early childhood, he had certain typical traumatic experiences, then he would now have these symptoms; he has the symptoms; therefore, he had the early traumatic experiences. If the patient does not recall such experiences, then the psychoanalyst says that he is resisting the therapy, and puts him through something like a third-degree examination until he confesses or quits.

In summary: to argue to the truth of an hypothesis by affirming the consequents is always a fallacy no matter how many consequents are affirmed. We are at liberty to entertain such hypotheses as have not been falsified by the observed failure of their consequents, but we may discard them whenever we find others that are more fertile or more simplifying. Inductive reasoning does not establish the *truth* of its hypotheses. Even if the hypothesis is false, half of its implicates will be true. But to say this is not to deny the value of hypotheses. They reveal interconnections between events, and through these interconnections, they suggest experimental procedures by means of which implicates can be verified or disverified. The verification or disverification of the implicates is important and useful in its own account, for by it, we discover something about the world in which we live, but the verification of the implicates does not establish the truth of the hypothesis.

HARRY M. JOHNSON

RED LIGHT AND THE FUNCTION OF THE RODS

* The careful study by Griffin, Hubbard, and Wald of human visual responses to long-wave stimulation makes it abundantly clear that foveal and peripheral thresholds for light beyond about 650 mμ are effectively identical.¹ The idea nevertheless persists that rods are insensitive to the extreme red end of the spectrum. After Dr. Herschel Leibowitz called this situation to my attention in a personal communication, I undertook to trace the erroneous notion to its roots.

¹ D. R. Griffin, R. Hubbard, and G. Wald, The sensitivity of the human eye to infra-red radiation. *J. Opt. Soc. Amer.*, 37, 1947, 546-554.

There are two principal empirical findings to note. First, extreme red light is the best of all possible pre-adapting stimuli for dark-adaptation. Because dark-adaptation proceeds so rapidly even after prolonged stimulation by red light bright enough to read by, it is easy to understand why visual workers have tended to conclude that the rods are not affected or not much affected by red light. Actually, red light merely produces less light-adaptation than other wave-lengths do; but it affects the rods as much as it does the cones. The effect is merely a relative one.

Secondly, if dark-adaptation is traced with a pure long-wave light, its course is fairly rapid and the final threshold much higher than for short-wave or for white light. Even when extra-foveal regions are involved, the complete adaptation-curve looks strikingly like one in which stimulation has been confined to the fovea. Again, the assumption that only cones are stimulated by such light appears pardonable. Instead, the truth seems to be merely that lights of different wave-lengths produce different degrees of adaptation and that in the function of adaptation (as in luminosity and threshold functions) red light affects all parts of the retina approximately alike while other monochromatic lights affect periphery and fovea differentially.

This situation is certainly not made clear in current secondary sources. For example, Morgan and Stellar write:

If the stimulus is in the extreme red region *which is covered only by the photopic visibility curve*, only the cone portion of the curve appears.²

Geldard, in his recent work, puts it as follows:

With extreme red light, *to which the rods are quite insensitive*, only the cone portion appears.³

Osgood does not plump quite so clearly for an inactivity of the rods, but nevertheless states:

Testing with extreme monochromatic red light yields *a pure cone adaptation curve*.⁴

Woodworth and Schlosberg come closer to a satisfactory presentation but still run the risk of leaving the reader with the old erroneous view. They write (referring to a figure of Judd's that presents threshold data correctly):

Reference to Figure 13.3 will show that red light is a reasonably effective stimulus for cones. If a man wears special red goggles that transmit only this band, he can

² C. T. Morgan, and E. Stellar, *Physiological Psychology*, 2nd ed., 1950, 160. (Italics in this and succeeding quotations are ours.)

³ F. A. Geldard, *The Human Senses*, 1953, 36.

⁴ E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 160.

read with his cones, while the rods are *relatively unstimulated* by the room light and are adapting.⁵

'Relatively' should mean relative to their own maximal sensitivity; but it is fair to guess that many readers will come away with the impression that it means relative to the cones, an unfounded view. Bartley is better, but still not emphatic enough to make the point.

Were we to use exposure to red light as provided by a filter with a cutoff at about 620 mμ (see Figure) we should be exposing the two classes of receptors (rods and cones) to different fractions of the total spectral ranges to which they are sensitive. This is one of the principal factors involved when red filters are used. To illustrate, we may again refer to the two luminosity curves (see Figure) where the portions of the two curves eliminated by the red filter are shown by the unshaded areas. It will be seen that a greater portion of the scotopic curve (red) is eliminated than is taken from the photopic curve (cone). The red filter passes 1/100 of the effective flux for the rods and 1/10 for the cones. This is a 10 to 1 difference, and it means that by wearing the red filter goggles, the individual reduces the effective flux of light to the rods more than to the cones.⁶

My only quarrel with this otherwise admirable exposition is that it should perhaps have concluded with "the individual reduces the effective flux relative to maximal rod response more than to maximal cone response." A casual or hasty reader may still take Bartley's summarizing sentence to mean that the red light has (absolutely) less stimulating effectiveness for rods than for cones.

The only author I have found whose comments are perfectly explicit on this matter is Chapanis who (in writing of the fact that readers often do not notice that luminosity curves are typically drawn relative to their own maxima) phrases it this way:

This misunderstanding leads to the frequent mis-statement that the rods are not as sensitive to red as the cones. Several lines of evidence, in particular recent experiments by Chapanis and by Wald, show conclusively that the rods are, in fact, just as sensitive to red as are the cones.⁷

Where did the "frequent mis-statement," persistent as it is, come from in the first place? That is more difficult to discover. Certainly Hecht is partially responsible. In his well-known 1934 article he discussed the pure-cone adaptation-function of night-blind patients, and then continued:

A similar situation exists in the measurements made by me (1921 and 1922) and by Kohlrausch (1922 and 1931) when adaptation is measured with extreme monochromatic red light to which at threshold *the rods are practically insensitive*.⁸

⁵ R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1954, 370.

⁶ S. H. Bartley, The psychophysiology of vision, in S. S. Stevens, *Handbook of Experimental Psychology*, 1951, 948.

⁷ Alphonse Chapanis, How we see, in National Research Council *Human Factors in Undersea Warfare*, 1949, 12-13.

⁸ Selig Hecht, Vision: II. The nature of the photoreceptor process, in C. Murchison, *Handbook of General Experimental Psychology*, 1934, 727.

Note the similarity of phrasing between this and some of the later accounts quoted above. In the original report of the 1921-1922 work to which Hecht refers, he had put it this way:

On the side facing the observer is a Wratten filter No. 70. This limits the transmission to the rays beyond 650 $\mu\mu$, and thus *insures participation of the cones only*.⁹

Kohlrausch, whose work is quoted with approval by Hecht and by Chapanis in different connections, plays a curious rôle. As Chapanis points out in the chapter already quoted, Kohlrausch recognized that the important fact for adaptation was the *relative* effectiveness of the light in question for rods and cones. On the other hand, in one publication he puts the effectiveness for rods of light beyond 680 $m\mu$ at *zero*.¹⁰ This he deduces from the fact that the photochromatic interval is zero, a fact entirely in agreement with modern findings!

In sum, the persistent notion that red light from the extreme long-wave end of the spectrum is effective for cones but not for rods seems to have had its origin in a psychological assumption about the duplex nature of the retina. That assumption ran about as follows: monochromatic red light from the extreme end of the spectrum produces the same experience in fovea and periphery; other lights show differential effects, either in chromatic threshold, apparent hue, or both; since the dual effects are attributed to stimulation of two types of receptor, single effects must mean only one receptor is affected. Now the conclusion is of course a *non-sequitur*. It appears, nevertheless, to have provided a basis for more than thirty years of error.

Swarthmore College

W. C. H. PRENTICE

SIXTY-SECOND ANNUAL MEETING OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

The Sixty-second Annual Meeting of the American Psychological Association was held in New York City at the Hotels Statler and New Yorker, September 2-7, 1954. In attendance were approximately 6,000 psychologists and a few friends. Each of the 17 divisions of the Association staged its program, resulting in the presentation of a total of 445 research papers and

⁹ Hecht, The nature of foveal dark adaptation, *J. Gen. Physiol.*, 4, 1921-22, 113-139, 116.

¹⁰ A. Kohlrausch, Tagessehen, Dämmersehen, Adaptation, in A. Bethe (ed.), *Handbuch der normalen und pathologischen Physiologie*, 2, 1931, footnote pp. 375f.

86 symposia. Though in both attendance and complexity the meetings were unique, they were, through the good and vigorous efforts of the Local Arrangements Committee, remarkably well organized.

Dr. O. Hobart Mowrer's presidential address, entitled "The psychologist looks at language," presented a theoretical analysis of linguistic behavior in terms of conditioned response psychology.

In its business meeting, the APA crystallized a number of actions that will affect psychology as science and as profession. Many will regard as the most significant action of 1954 the creation of the Scientific Development Board—a Board instructed to define and forward activities by which the resources of the Association can functionally be turned to the advancement of psychology as science. Another significant action was the formation of the separately incorporated American Board of Psychological Services, an organization whose function will be to publish a directory giving to the general public a list of agencies and individuals properly qualified to render personal psychological services.

During the meetings it was announced that Theodore M. Newcomb had been elected President-elect, that Ann Magaret Garner and D. O. Hebb had been elected to the Board of Directors, and that M. Brewster Smith would be the new Editor of the *Journal of Abnormal and Social Psychology*, with his term beginning January, 1955.

In 1955, the Annual Meetings will be in San Francisco, in 1956, in Chicago, and in 1957, probably in Miami Beach.

Washington, D.C.

FILLMORE H. SANFORD

Walter Samuel Hunter: 1889-1954

Walter S. Hunter, one of the world's most productive leaders in biologically based scientific psychology, died suddenly on August 3, 1954, in Providence, Rhode Island. His branch of the Hunter family came to this continent from Scotland via the north of Ireland in the eighteenth century. His immediate forebears moved to the Ohio country soon after the Revolution. He was born on March 22, 1889, in Decatur, Illinois. He received his bachelor's degree from the University of Texas in 1910. His first plan had been to study electrical engineering, but he became interested in psychology and, while still an undergraduate and due to the influence of C. S. Yoakum, he conducted an experimental study in animal behavior. This investigation was later published as "Some labyrinth habits of the domestic pigeon."¹

His graduate work was done at the University of Chicago where he was influenced especially by Harvey Carr and James Roland Angell. His doctor's dissertation, "Delayed reaction in animals and children" is regarded as a classic in its field.²

He taught first at Texas and then at twenty-seven years of age became head of the Department of Psychology at the University of Kansas. In 1925 he became G. Stanley Hall Professor of Genetic Psychology at Clark University. In 1936 he became head of the Département and director of the Psychological Laboratory at Brown University. He was active in many general University affairs at Brown. This spring on his sixty-fifth birthday he resigned the chairmanship of the Brown Department, but he had looked forward to at least five more years as an active investigator and teacher before retirement. After the death of his first wife, Katharine Pratt, in 1915, he married, in 1917, Alda Barber, herself a student of comparative psychology. Mrs. Hunter and two daughters survive him.

Hunter's greatness as a pioneer and as a leader of his generation in the scientific study of behavior lay in his ability to see the mammal and man as total behaving systems existing in both physical and social environments. He was a psychologist with an interest in the nervous system and the receptor processes but this did not make him forget that the explanation of human behavior is the main aim of scientific psychology. His keen mathe-

¹ *Op. cit.*, *J. Anim. Behav.*, 1, 1911, 278-304.

² *Op. cit.*, *Behav. Monog.*, 2, 1913, (no. 6), 1-86.

mathematical mind led him to quantify in new ways relationships that had previously been dealt with in a qualitative and descriptive manner. This tendency is to be seen in his work with the delayed reaction and in his study of the reliability of the maze and other learning problems. His interpretation of the phenomenon of retinal streaming in relation to the after-effect of visual movement and his interest in the span of visual discrimination in relation to the Bunsen-Roscoe law and many other studies attest the effectiveness of this attitude. His broad approach to behavior is demonstrated in his early study of American Indians of mixed ancestry in a consideration of the relationship between genetic background and intelligence.

He served his professional colleagues in many ways. He was at various times concerned in editing the *Journal of Animal Behavior*, the *Psychological Bulletin*, *Behavior Monographs*, and *Comparative Psychology Monographs*. For the past fourteen years he was an associate editor of this JOURNAL. He contributed four chapters to books published by the Clark University Press. His basic textbooks, *General Psychology* which went through two editions, the second of which was translated into Chinese,³ and *Human Behavior*,⁴ were clearly and concisely written. They had an important influence upon many mature psychologists as well as beginning students. He won wide recognition for a series of theoretical papers in which he advocated the use of the word *anthroponomy* as an objective name for the field commonly called scientific psychology. In connection with this proposal he has said, "I was never under any delusion that the designation of the science would be changed to anthroponomy, but the paths of our science would have been much smoother in its public relations had some nonpsychic term designated it."⁵

In 1925, he became the editor of the *Psychological Index*, and from its beginning in 1927 until 1946 he was the editor and indeed the creator of that great periodical tool of all his professional colleagues, *The Psychological Abstracts*. In connection with this work he came to have a wide acquaintance among European and other non-American psychologists.

He served as chairman of the Division of Anthropology and Psychology of the National Research Council from 1936 to 1938. In the First World War, as a commissioned officer, he was chief psychological examiner at Camps Lee, Sheridan, and Devens. Characteristically, some of the work that he did at this time was fundamental in the later acceptance of the whole

³ *Op. cit.*, 1st ed., 1919, 1-351; 2nd ed., 1923, 1-368.; Chinese translation of second edition in 1926.

⁴ *Op. cit.*, 1928, 1-355. Completely revised edition of his *General Psychology*, with change of title.

⁵ *A History of Psychology in Autobiography*, 4, 1952, 172-187.

idea of psychological examining by American military authorities.

During the Second World War he served as a member of the Emergency Committee in Psychology and later as chief of the Applied Psychology Panel of the National Defense Research Committee. For this work he was given the President's Medal of Merit. Because of his broad psychological knowledge and his ability to make sound human judgments he was called upon to serve as a consultant and as an adviser to many special committees associated with philanthropic foundations and with national and international scientific associations as well as with various branches of the Federal Government. He was president of the American Psychological Association (1931), the Eastern Psychological Association (1940-41) and a member of and active in the management of both the National Academy of Sciences (1935 *et seq.*) and the American Philosophical Society (1941 *et seq.*). One of his avocations was painting and some of his canvases are most pleasing.

Walter Hunter had a deep and enduring interest in the basic problems of human nature. His colleagues always knew him as a friendly fellow human being. He had that rare ability given to few psychologists of using his own understanding of animal and physiological psychology in the solution of the living problems of everyday human life. He was scientifically learned and humanly wise.

The Smithsonian Institution

LEONARD CARMICHAEL

AN ACKNOWLEDGMENT

The JOURNAL is indebted to Mrs. Walter S. Hunter for the photograph reproduced in the frontispiece of this number. It was taken by Fabian Bachrach of Providence, Rhode Island, in 1950 when Professor Hunter was 61 years old. The signature, a facsimile of which appears under the portrait, was taken from a recent letter.

K. M. D.

BOOK REVIEWS

Edited by M. E. BITTERMAN, University of California

Physiological Acoustics. By ERNEST GLEN WEVER and MERLE LAWRENCE. Princeton University Press, 1954. Pp. xii, 454.

Wever has been most successful in auditory research for over twenty-five years, and the appearance of this comprehensive volume will be welcomed by every worker in the field. The authors define "physiological acoustics" as concerned with those parts of auditory perception in which the function of the ear is chiefly acoustical, and they thereby exclude all neurological and electrophysiological considerations. The restriction thus imposed is essentially adhered to, except in certain portions of Chapters 14 and 18. Subjective aspects of hearing are not systematically treated, but necessary psychological data are included in the chapters on sensitivity, distortion, and auditory theory. Concentrating their effort on the peripheral part of auditory function, the authors are able to give a thorough account of a field in which the contributions of Wever and his collaborators have been most prominent. The book also contains some pertinent experiments which are as yet unpublished, a fact which raises its value above that of a merely compilatory study.

The introductory part of the book begins with a brief historical outline and a survey of anatomical facts. An elementary account of basic acoustics, including measurement of sound-intensity, units, and reference levels, follows. The next chapter describes experimental methods used in physiological studies of the peripheral ear. It deals predominantly with procedures employed in the authors' laboratories. Although the account covers only sinusoidal stimuli and omits special recording methods as used by Békésy, Tasaki, and others, it contains many details that will be useful to the specialist. The final section of the introductory part deals with experimental methods for determining threshold sensitivity, conditions affecting these measurements, and the determination of the mechanical impedance of the ear. Returning to the problem of threshold, the authors conclude that thermal motion of air molecules should not limit auditory sensitivity, but that physiological conditions, such as muscular contractions, vascular noise, metabolic processes, and the like, may be important limiting factors.

The second part of the book deals in detail with the action of the middle-ear apparatus as a mechanical transformer required to minimize loss of energy in the transition from aerial to liquid vibrations. As a result of a combination of original studies with those of others, it is concluded that the ratio between the areas of drum membrane and stapes footplate, as well as ossicular lever-ratio, are important for the effect of the middle-ear transformer. The impedance transformation is estimated for the ear of cat and man.

The third part of the book, dealing with nature and locus of distortion, is a summary and extension of the important original work that Wever and his collaborators have done in the field. Frequency and phase distortion are discussed, and their measurement by microphonic methods are described in detail. An account of the indirect determination of distortion caused by the stapedial function is given.

Quantitative aspects of amplitude-distortion, including measurements of harmonics and combination tones, are carefully explained. A summary of the experiments demonstrates that amplitude-distortion occurs mainly in the inner ear. Further analysis shows that microphonic distortion is probably connected with electro-mechanical processes and not with the more peripheral events of sound conduction through stapes footplate and inner-ear fluids. Finally, an analysis of cochlear interference (partial suppression of the microphonic response by simultaneous application of a second stimulus) and other results leads the authors to the conclusion that harmonic distortion and cochlear interference occur earlier in the process of hearing than does overloading.

The chapters which comprise the fourth part of the volume deal with the function of the middle-ear muscles and the functional effects of a change of air-pressure in the external and middle ear.

The fifth part of the book, entitled "Sound Conduction in the Cochlea," contains more than its title indicates, dealing with the following four problems: (1) the pathways of entry of sound to the cochlea; (2) the manner of sound transmission within the cochlea; (3) the spatial distribution of cochlear activity; and (4) the nature of the local stimulating process. In the section on pathways to the cochlea, air-conduction is first treated. Then sound-transmission through the round window, a problem of long-standing interest for the otologist, is discussed, and found to be of negligible importance in the normal ear. The lucid treatment of bone-conduction will be of great interest because it stresses the importance of a correct formulation of this difficult problem. On the basis of experimental evidence obtained by the authors and others, it is concluded that in bone-conduction the translatory mode of vibration is of greatest importance. This means that, due to inertial lag, a vibration of the cochlear capsule results in a motion of the cochlear contents quite similar to the effect of air-conducted sound. It can also be shown that a stimulus involves the same sensory cells, no matter if it is applied by air or bone conduction. A study of the functional effects of blocking of the oval or round windows closes the chapter on sound-conduction through the cochlea. These last two problems are important for clinical application.

The question of sound-conduction within the cochlea leads into the problematic field of auditory theory. A brief discussion of the place principle, which explains pitch-discrimination on the basis of stimulation of specific loci along the basilar membrane, prepares the reader for the discussion of hearing theories. Resonance theories essentially depend on mechanical properties of specific areas along the basilar membrane which are supposed to be stimulated as individual resonators. The chief difficulty of resonance theories is that a resonator can either be sharply tuned or strongly damped, but not both. Since elementary observation shows very considerable damping in the function of our ear, localization on the basis of resonance alone would have to be very inaccurate. Resonance theories do not require specific assumptions concerning the speed of wave-propagation. A certain degree of specificity is also obtainable by assuming that the wave set up in the inner ear is propagated along the basilar structures with an amplitude reaching its maximum in different areas according to frequency. This assumption can be justified on the basis of hydrodynamical principles, considering the elastic properties of the wall of the vibratory fluid. These "traveling-wave theories" do not depend on the resonance principle. The more recent theories of this type (and of the standing-wave type) are

discussed with great skill, their original mathematical form being omitted for the convenience of the reader. In a general evaluation of all these theories, the authors try to point out assumptions common to all of them and show that slow-wave propagation is an essential feature. There exists a close relation between conductive speed and lowest discriminable pitch which is tabulated for the more important traveling-wave theories. In a critical discussion, the authors produce experimental evidence, based on cochlear microphonics, that an estimate of the speed of conduction fails to agree with the slow propagation of sound within the cochlea as postulated by these theories. Experimental results, seemingly supporting traveling-wave theories, as obtained by Békésy, Tasaki, and others, are discussed but not considered to be decisive.

The first part of the section on cochlear patterns and the stimulating process contains some details concerning the shape and extent of stimulation demonstrable along the basilar structures. Based on combined evidence from histology following stimulation-deafness and microphonic response, including its essentially linear intensity function, the conclusion is reached that every tone applied spreads its effect throughout the cochlea although the patterns differ for different tones, being more compact for higher frequencies. Details of the mechanical effect exerted on the hair cells by sound are still hypothetical. Various hypotheses concerning this problem are discussed. They include several attempts to introduce mechanisms which would tend to render the area stimulated by pure tones more circumscribed.

The chapter concludes with a discussion of the origin of cochlear potentials and the excitation of the auditory nerve. These two sections fall, strictly speaking, outside the plan of the book since they contain electro-physiological material. The reader will, however, welcome their inclusion since they contain a brief account of Békésy's work on the *DC* potential within the inner ear. Concerning the initiation of nerve-stimulation, the authors defend the standpoint that cochlear potentials directly cause neural activity. The action of a chemical mediator is discussed, but experimental evidence in favor of this hypothesis is considered insufficient.

The sixth part of the book constitutes a clinical interlude. The authors have been closely connected with Dr. Lempert and his institute for several years, receiving stimulation for the study of problems in auditory pathology in general, and the problem of otosclerosis in particular. Several important papers on these subjects have been published by Wever and his collaborators, and it appears natural that this volume, which is dedicated to Dr. Lempert, should contain a section dealing with clinical material. There is first a chapter on conductive deafness discussing differential diagnosis, tuning-fork tests, and types of hearing-impairment following pathological or experimental interference with various parts of the sound-transmitting mechanism. Another chapter deals with various aspects of otosclerosis, a progressive type of deafness caused by increasing fixation of the stapes due to bony growth. After a discussion of the pathology of the disease, the fenestration operation (creating a new opening in the labyrinth) is described and its success in improving hearing demonstrated. The rest of the chapter is devoted to various aspects of the physiology of the otosclerotic and of the fenestrated ear. It is pointed out that due to loss of the transformer-action of the middle ear, the fenestrated ear cannot have normal hearing. A loss of 25 to 30 db. (which, however, does not constitute a serious handicap) is to be expected.

In the final part of the book, there is a return to physiological considerations.

The first section treats the problem of acoustical impedance of middle and inner ear as a function of frequency. These results depend upon rather difficult and indirect evidence obtained by a number of authors. The mechanical properties of the middle-ear apparatus are further analyzed in the subsequent section. It is pointed out that the various mathematical models described so far do not permit satisfactory predictions about the actual behavior of the middle ear since important empirical data are still missing. The authors then proceed to their own analysis of the sound-conducting apparatus, including the stapes. Using response-characteristics obtained by microphonic intensity and phase measurements, they construct an electrical circuit and its mechanical equivalent which will display the same resonance characteristics as the middle ear of the cat. Since numerical values cannot be assigned so that predictions of mechanical behavior can only be qualitative.

The final section returns to the problem of auditory theory, the subject of Weaver's last major work. There is a restatement of the difficulties which any explanation of pitch-discrimination based on a principle of narrow localization has to overcome: all experimental evidence points to a rather diffuse representation of pitch along the basilar membrane, whereas pitch-discrimination is remarkably sharp. "Sharpening up" processes as advocated by several authors are briefly discussed. It is pointed out that frequency-localization at higher neural levels, as determined by electro-physiological methods, also lacks the precision of our subjective auditory experience. As in Wever's former book, the "frequency principle," according to which neural impulses are synchronized to the stimulating frequency, is introduced. It is maintained that together the place principle (for the highest tones) and the frequency principle (for low tones) or a combination of the two (for the intermediate range) can account satisfactorily for the wealth of our auditory experience.

The book contains several valuable appendices, including a comparison of reference-levels for sound intensity, a comprehensive list of measurements of auditory structures, and a decibel table with instructions for its use. The index and list of references are very comprehensive. The illustrative material and diagrams are excellent. Very few typographical errors were encountered. In line 23 of page 242 "round window" should be replaced by "oval window." In the caption of Figures 106 and 107, the reference should read "Wever and Lawrence (8)" instead of "Wever and Lawrence (9)."

If the book is to be characterized by a few words, it would seem to be a condensation of work and thought of one of the most prominent scientists in the field of audition and of his school. It will be profitably consulted and studied by all people actively engaged in auditory research, including some otologists. It is also an intellectual treat: a systematic presentation of carefully obtained experimental evidence, followed by penetrating and unrelenting analysis, sometimes reminiscent of the procedures of a competent defense lawyer, amply reward the careful reader.

The work does not compete with handbook articles. It presents in many aspects the opinion of Dr. Wever and his school, a reminder of the vastness and importance of their past work, which is also expressed by the material covered in the index. There one finds 45 references to Békésy, 6 to Davis, 45 to Lawrence, and 77 to Wever. The mature reader will undoubtedly find many statements challenging and possibly be stimulated to do further work in the field.

The part of the book concerned with clinical applications, presumably written with otological assistance, may merit some minor comments for consideration in

a later edition. In the section on differential diagnosis of types of deafness it is stated that scarlet fever produces damage mainly to cochlear and neural structures. This may not be a generally valid contention. The statement that a person with sensory or neural deafness is commonly unable to use a hearing aid to advantage appears too strong and untenable in view of general audiological experience. The same chapter does not mention the phenomenon of recruitment, which is one of the more important tools in differential diagnosis. On page 354 we read: "It is evident that the hearing in such cases might be improved if something could be done to increase further the difference in stimulus of the two windows." Here, experimental and clinical studies of the artificial drum ought to be included.

The position on hearing theory coincides with the one taken in Wever's previous book. The authors appear very successful in pointing out deficiencies of place theories. To account for the precision of our pitch experience, additional neural help seems to be inevitable. The combination of place and frequency principles introduces an element of dualism which will probably be offensive to many, but it seems to be one of the lessons of modern science that theories should not be judged on the basis of esthetic principles. Actually, recent experimental evidence not treated in the book indicates that the perception of pitch depends on the combined action of the peripheral and central auditory mechanisms. Whether the latter can be accounted for by the simple assumption of synchronization of nervous impulses, remains a problem for future research. It is the opinion of the reviewer that the final answer will be more complicated.

KARL LOWY

University of Rochester

Principles of Industrial Psychology. By THOMAS A. RYAN and PATRICIA C. SMITH. New York, Ronald Press, 1954. Pp. xiv, 553.

Industrial psychology has attained stature and breadth since Hugo Munsterberg pioneered the field forty years ago. It has yet to attain a coherent, interlocking set of principles that govern the topics encompassed within the term. A purist, therefore, might quarrel with the title of this book and insist that between the covers we can only find some currently useful but unconnected generalizations about the psychological problems of industry. The field lacks some unifying psychological concepts that bring together selection, training, and problems of production. Economically, the industrial enterprise is unified by concepts centered on maximized profit. Psychologically we need some industrial principles whose hub is the individual, his growth, his value to himself and to his working group. To those in the field such a criticism is old hat, yet we need to be reminded from time to time of these limitations, even if the time is not ripe, or our insights are too blurred to achieve synthesis.

This book should be proof, if any is needed, that psychologists have made important and telling contributions to industrial operations. Nevertheless psychologists are more often than not still regarded by many business executives as too academic and unrealistic. The psychologists on their part have been appalled by the business man's scientific naïvete in matters of human behavior. Ryan's and Smith's book will in all probability reinforce the business man's bias (if he reads it). On the other hand, it should provide the undergraduate, who may later become a business executive, with an appreciation of the scientific approach to some human problems of industry.

Aside from the introductory chapter which presents a description of the field and the activities of an industrial psychologist, the remaining 18 chapters are divided into two sections—one on Selection and Placement and the other on Factors in Efficiency. Ryan and Smith see the industrial psychologist as essentially a research individual gathering facts and recommending action, but taking almost no part in administration. In fact, the authors specifically eliminate administrative skills from consideration. This position provokes a question as to why psychologists have shied away from the social psychology of administration and why, even though the dependable data are sparse, should not at least one chapter be devoted to these significant problems? We have not hesitated to prescribe parent-child relationships. We have studied the social development of children and adolescent groups. Why not explore some research problems in administration just as the authors have done within the fields they stake out as the legitimate areas of industrial psychology? Perhaps if we tackled these problems our approach to some of the other aspects would be less subject to the 'ivory tower' criticism.

The complications that stem from a non-administrative rôle for industrial psychology is evidenced in the chapters dealing with selection and placement of employees. Administrative operations of selection and placement programs in live, vibrant organizations intrude themselves at various points in a way that gives evidence that Ryan and Smith have indeed experienced these restrictions on their otherwise uncontaminated research efforts. They write with a sense of practical realities.

Large-scale industrial organizations never confuse their research laboratories with their factories. Both have a set of governing principles and are related to each other in a rational way. As industrial psychologists we have not always kept these distinctions clear nor have we made explicit how research results and procedures need to be modified for operations. To a degree this confusion runs through the first part of the book, although the final chapter of Part I describes the operation of a selection and placement office.

From the standpoint of coverage and wealth of references, Part I leaves little to be desired. A chapter is devoted to each of the following topics: Principles of Selection, Job Analysis, Validation of Predictors, Interviewing and other Miscellaneous Methods of Selecting, Intelligence Testing, Aptitude Tests, Interest and Personality Scales, Methods of Placement. The authors have obviously struggled to organize their material in a sensible sequence. It is always difficult to know just how to weave in the necessary statistical background in presenting the problems of selection. Ryan and Smith have wisely chosen to discuss the statistical concepts as they are necessary for an understanding of the procedures, rather than devote a separate chapter or section to statistics in the abstract. Moreover they have minimized formulas and their derivations, in favor of a commendable exposition of the logic of statistical treatment. One can always find details not to his liking. Figure 7 is a misleading scatter-diagram that no statistician would ever construct, and the fact that it was modified from another source does not temper the criticism. The discussion of job analysis is scattered over 20 pages with a number of digressions interrupting the main theme. These are perhaps only the private complaints of a lone reader.

A very commendable feature of Part I is the use of a particular case as a con-

necting thread between the chapters. The case describes the various steps taken and choices made by a psychologist from the initial consulting invitation to the final validation. Chapter 9 dealing with placement of employees within an organization is a unique feature of the book which may also become a controversial topic. The business executives will probably look upon it as unrealistic, although with greater or less success the basic notions described here have been employed by the armed forces.

Certain features of Part II demand special comment—both favorable and otherwise. Chapter 11, "The Meaning of Human Efficiency," is a thoughtful and careful effort to pack together the various operational meanings of work-cost with a minimum of descriptive detail pertaining to specific studies by each means. The presentation here has the teaching advantage of getting the various meanings of effort into one area for purposes of comparison. In most other texts these concepts are scattered and separated by other topics.

Three chapters are devoted to industrial motivation of which the first two are largely theoretical. These carry the reader to a higher level of abstraction and theoretical sophistication than any other part of the book. A change of pace midway in a course may have its advantages. For some students the more abstract material may undoubtedly be welcome, and the effort may well be a step in the direction of the synthesis which the whole field needs. All things considered, however, these chapters seem out of place, valuable as they may be for a monograph or journal articles.

Somewhat later—in the chapter on accidents—the authors again use the pages of the book to debate the concept of accident proneness in much the fashion of a controversy that might better be carried on in a journal. This is not to deprecate raising bona fide issues in a text. In the interest of balance, every textbook author must exercise a certain restraint in writing on those topics on which he may have special knowledge. Ryan and Smith seem to have let themselves go on these points with the result that it is difficult to know whether the book is best placed early or late in the curriculum. The authors themselves do not specify for what level they were writing.

A single chapter—albeit a long one—is given to learning in industry with major emphasis on methods of facilitating the process. This chapter represents one of the better ones in the entire book drawing upon a broad time range of studies and written with a confident, non-querulous tone. Perhaps the confidence of this chapter is a reflection of the maturity of research in this field and the fact that many of the researchable issues have been debated into useful shape.

If there is any unifying concept that pervades industrial psychology, it is a point of view and a method of study. The unique value of psychology to manufacturing and associated activities lies in the psychologist's special talents for experimenting with human beings and researching human problems. Ryan and Smith's book will certainly carry that message, well documented with studies ranging from those recently published, to familiar old-timers, which have appeared in numerous earlier textbooks. *Principles of Industrial Psychology* does not have the smoke and drive of the factory—but it does convey the impression of dignity, maturity, and confidence of a profession that knows what it is doing.

Institute for Research in Human Relations

F. K. BERRIEN

The Nature of Prejudice. By GORDON W. ALLPORT. Cambridge, Addison-Wesley, 1954. Pp. xviii, 537.

An evaluation of Professor Allport's latest book is complicated by the fact that it is written by a gentleman scholar about an ungentlemanly topic for a non-scholarly audience. It is directed explicitly toward two groups—students and general readers concerned with attaining a better understanding of prejudice.

Viewed on this level, the volume should be of considerable value. It is highly readable—one has a lurking suspicion that the manuscript was gone over with a Flesch count in hand—and only an occasional polysyllabic word betrays the author. Points are effectively illustrated with anecdotes, many of which are amusing as well as pertinent. As a general introduction to current scientific thinking about prejudice it is most satisfactory. There are noteworthy examples of reducing complex matters to a level which should be readily understood by the general reader. The chapter on "The Scientific Study of Group Differences" is especially impressive.

Psychologists also will find much of interest. The emphasis upon the multi-determined nature of prejudice is a welcome antidote to narrowly formulated approaches which stress one class of presumed causal factors to the exclusion of others. In casting his net wide the author has come up with much fascinating and little known, pertinent data. The scope of the work is indicated by the fact that 31 chapters were required—systematically progressing from "preferential thinking" through group differences, perceiving and thinking about group differences, socio-cultural factors, acquiring prejudice, the dynamics of prejudice, and character structure, to reducing group tensions. Such breadth has concomitant disadvantages since specific topics are necessarily simplified in the selection and summarization of relevant data (the chapters average less than 17 pages in length). Specialists in the areas touched upon may rightly feel that some generalizations do not do full justice to the richness and complexity of existing knowledge.

A related problem concerns the interpretation of research findings. In general, Allport's interpretations would be concurred in by most scholars and their validity is suggested by the general consistency of results from many studies. Nevertheless, the reviewer felt occasional twinges of disquietude at some summaries of research findings. On p. 403, for example, Reichard's cautious comparison of Rorschach protocols of college women high and low in prejudice is interpreted as indicating inhibition and compulsive overmeticulousness on the part of prejudiced Ss, whereas those with less prejudice gave more whole responses and had more productive associations. It is not noted, however, (1) that these findings were not statistically significant (only one of 49 tested relationships reached the .02 level of confidence), (2) that many of the differences found are related to I.Q. (which was uncontrolled), (3) that many hypothesized differences between those high and low in prejudice were not confirmed, and (4) that the differentiation of the two groups on responses to the Rorschach was *ad hoc*. Indirect evidence from other sources suggests that most of Reichard's findings could probably be replicated but, as far as the reviewer knows, no one has done so.

Professor Allport cannot be held accountable for methodological limitations of much research on prejudice. It is not the proper function of such a general survey of research as he has undertaken to go into detail about the limitations and possible methodological weaknesses of much of the relevant research. His account is not

likely to mislead seriously the audience to which it is directed; indeed, it is difficult to imagine a better summary. His book may, however, give the impression that knowledge about prejudice is more unequivocal and rigorously based than it actually is.

The serious student can use this book as a general summary of the best informed thinking on prejudice and as a guide to the research literature. There are other good reasons for examining this book carefully: (1) its breadth insures that certain topics will be touched upon which are unfamiliar to many concerned with the general area; (2) a good deal of unpublished research is cited which is a major addition to the body of available information; (3) many of Allport's tentative generalizations based upon admittedly inadequate data suggest hypotheses for research (the chapter on "Social Structure and Cultural Patterns" is especially provocative in this respect). One is tempted to add a fourth reason. Those who write about prejudice would be well advised to model the tenor of their remarks upon the restrained example set by the author.

The reviewer agrees with Professor Allport that this account of the nature of prejudice may soon be dated. He is further convinced that the more widely the book is read and understood (by both lay and scientific audiences), the more rapid will be the dating of the material presented.

RICHARD CHRISTIE

Research Center for Human Relations,
New York University

Learning Theory, Personality Theory, and Clinical Research: The Kentucky Symposium. By VARIOUS AUTHORS. New York, John Wiley and Sons, 1954. Pp. ix, 164.

This volume preserves for posterity a group of eleven papers read last year in a symposium sponsored by the University of Kentucky to encourage "a closer integration" of the "three branches of psychology" named in the title. A large amount of material only indirectly related to that laudable purpose managed to find its way into the papers, but each contributor got around sooner or later to some opinion on the contribution of learning theory to clinical psychology and to the understanding of personality. It may be well to call the roll.

Mowrer, Spence, and Wickens are enthusiastic about the potentialities of learning theory, as is Ammons, who finds it necessary to speak in favor of theorizing in general (a "defensible procedure"). Adams, too, believes in the possibilities of synthesis, pinning his hopes on the unifying properties of the concept of *sentiment*—"the totality of conditions in the organism that make a given object what it is for the behaving animal." Harlow displays a certain amount of ambivalence, as does Cattell, who complains (oddly enough) that "learning theory has concerned itself with individual response variables instead of factor patterns and their measurement," but both are inclined to be friendly. Maier cautions against "premature crystallization" of narrow concepts. Snygg has nothing but contempt for the work that has been done in learning, rejecting the notion of learning as process. Wittenborn, in contrast, insists that we must study the "*how* of behavior change" rather than the "*content* or *what* of the change" which is but "an accident of time and place." He suggests, however, that "systematic and thoughtful observation of the changes in human behavior under diverse conditions can contribute more to learn-

ing theory than learning theory can at present contribute to the work of the clinical psychologist." In what is perhaps the most thoughtful and balanced attempt to deal with the central problem of the symposium, Butler sees learning theory as a means of combatting the "parochialism and intellectual provincialism" in contemporary therapeutic circles. If "theories of therapy and therapy-based personality theories are reduced to more general psychological theories which will 'contain' all of them," convergences can be noted and inconsistencies delineated. Such translations, of course, contribute nothing to our empirical knowledge of psychotherapy, and the danger must be recognized that "what are exact terms in learning theory then become as inexact as their semantic equivalents in psychotherapeutic and personality theory."

The volume contains many other interesting things. Mowrer decides to dispense with the concept of habit. Rewards and punishments, we are told, do not strengthen or weaken stimulus-response connections; "what they do instead is to produce, by conditioning, *secondary reinforcements* and *secondary motivations*, respectively"—"token decrements" and "token increments" in "emotional tension." The question of how these "tokens" function in response-selection is enveloped in a cybernetical cloud. Spence, on the other hand, is more inclined than ever to avoid entirely the problem of what is learned and the conditions essential for learning. He writes: "Although the concept of S-R tendency will be used in the ensuing presentation, no bias is intended as to whether the hypothetical learning change is a connection between two sensory elements of the nervous system or between afferent and efferent units. . . . Likewise the term 'reinforcement' will be used but only in the general sense that the stimulus effects or consequences of a response in some way or other are related to the changing value of the S-R tendency." Wickens uses the clinician's interest in personal determinants of perception to justify a lengthy and quite unsuccessful attempt to handle 'acquired distinctiveness' in terms of a watered-down version of the Spencian theory of two decades ago. The major part of Adams' paper is devoted to an abstruse discussion of the reality of psychical systems (principally directed, it seems, to the absent Krech). Ammons labors mightily with a set of postulates (somehow not quite as elegant as the Hullian set which supplied the vision) and brings forth the deduction (experimentally confirmed, as a matter of fact) that students will make more erasures and crossouts in the writing of a difficult examination than in the writing of an easy one. *Usw.*

We may suppose that everybody had a fine time at the symposium, and that for hosts and guests alike this printed record of the proceedings will provide a pleasant souvenir. For the rest of us, too, the enterprise may have some lasting value—as an indication of the curious state of psychology at mid-century.

M. E. B.

Meaning, Communication, and Value. By PAUL KECSKEMETI. Chicago, University of Chicago Press, 1952. Pp. viii, 349.

Kecskemeti, who received a Budapest Ph.D. in political science and sociology in 1928, is now employed by the Rand Corporation as a social scientist, but in this book he reveals himself as a philosopher. Whether the book will find wide acceptance among philosophers, I cannot say, but it appears to be a superlative piece of philosophical analysis. Because of the inherent nature of the work, as well as the

author's background in social science, it has intrinsic interest for psychologists.

In his preface (p. v), Kecskemeti states that his purpose is to show that interpretation and evaluation "inevitably include an element of *personal* judgment and decision," and that to renounce such subjective processes in favor of the "*impersonal* disciplines of formal, logical deduction and empirical fact-finding" would be to do a great disservice. Thus, the central theme is the treatment of meaning, defined (p. 1) as "anything that is or may be 'interpreted.'" Meaning is present only when some organism perceives a situation as presenting the possibility of some kind of "good" response. Standards of "goodness" include those which yield biological satisfaction, either directly or indirectly, and thus we are ultimately led, in the last section of the book, to a distinction between "relevance" (the meaning of a situation in terms of direct satisfaction) and "value" (the same, in terms of possible remote satisfactions), and, ultimately, to a consideration of utilitarian philosophies as providing a possible "autonomous" system of values. Kecskemeti's judgment on this latter point, incidentally, is in the negative; one of his arguments is that the utilitarian's complete ban on "moralizing" in child-rearing is psychologically unsound. Nevertheless, it might be objected, would not an enlightened (though hypothetical) utilitarianism permit "moralizing" if that happened to offer "utility"?

Along the way, the reader is offered intriguing discussions of a variety of topics. Memorable are: a new distinction between "signs" and "symbols" (26 ff.), an analysis of types and levels of messages (29 ff.), certain objections to logical formalism (71 ff.), the characterization of word-perception in reading as a "learned neural reflex" not subject to extinction (117), a distinction between "name-language" and "word-language" (125 ff.), the identification of word meanings as linguistic rules (137 ff.), a discussion of the functions of sentences, including an interesting definition of the "elementary sentence" (154), a listing of types of "value languages" (Chapt. VIII), a discussion of "levels of conflict" between values (246 ff.), a sketch of a philosophy of education ("education for success would be the greatest calamity that could befall our educational system," 267), thoughts on how a norm-oriented society can allow non-conformism (274 ff.), views on super-ego formation in the child (301 ff.), and finally, a way to validate judgments of value (Chapt. XI).

Kecskemeti feels (p. 90) that the less well-developed sciences (including, by clear implication, psychology) have need of the philosopher's guesses and criticisms. It is obvious that the author has read widely in psychology, yet he is pessimistic as to the ultimate possibilities of behavioral science. He doubts the possibility of a unified science, for he feels that organismic behavior displays non-mechanistic features, particularly where it displays meaningful responses. Description is possible, but explanation and prediction impossible. The argument is ingenious: if behavior can be predicted, "future scientists of behavior, when observing a conversation among people they do not understand, may nevertheless be able, under certain conditions, to determine what each of them will say." This implies the "paradoxical achievement" whereby the scientist is "able to 'catch up' with others having more information than he has himself" (p. 12). Kecskemeti neglects to point out that the science of behavior attempts only to determine relations between *given* antecedents and consequents; it can only specify possible "scenes" in the shape of things to come. Kecskemeti almost rules this out, however, when he decides (Chapt.

IV) that at least some behavior is truly free, and that behavior theory can only be probabilistic.

Whatever his philosophical predilections, the behavioral scientist will find this book rewarding. The predominant tone of pessimism with respect to a science of behavior is counterweighted by the careful analysis of the process of "interpretation."

Harvard University

JOHN B. CARROLL

Dynamic and Abnormal Psychology. By W. S. TAYLOR. New York, American Book Co., 1954. Pp. xiv, 658.

Among contemporary texts in abnormal psychology, Professor Taylor's book is unique. Despite its title, much of its content has little to do with either abnormal or dynamic psychology, as those areas usually are envisaged. One reason for this peculiarity may be that the author regards abnormal psychology as "the psychology of those mental states and of those minds which are relatively far from ideal integration" (p. 2), a definition which leads to such curious consequences as the classification of sleepiness and dreams as abnormal (even though not pathological). A further reason is that the author is primarily interested in the everyday, minor abnormalities of the various psychological functions of the individual, e.g. of perception, memory, or thought, and not in the more serious disorders which seem to involve the personality as a whole. The result of this attitude is that he devotes only 18 pages to the psychoneuroses, and only 27 to the psychoses.

Much of the book deals with the material of general psychology, and the titles of many of the chapters are identical with those of most texts in the general field, e.g. "Reactions," "Connector Processes," "Learning and Memory," "Acquired Patterns," "Attention and Preception," "Thought," and so forth, although—oddly enough—there is no chapter entitled "Personality." A student who already has had an adequate course in general psychology might find much of the book repetitious.

The manner of presenting each topic also is unusual, although it is reminiscent of that used by John J. B. Morgan in *The Psychology of Abnormal People* (Longmans, Green, 1936). Each major topic is divided into a number of sub-topics. In the case of learning these include conditioning, association, negative adaptation, experimental extinction, trial and error, immediate learning, insight, rote learning, meaningful learning, transfer, interference, automatization, higher units, reintegration, and the likes. Each sub-topic is allotted a brief description, followed by numerous illustrations of abnormalities in the function or activity in question. One result of this method is that the individual seems to be fractionated into a number of independent functions, and his behavior as a whole is not made understandable in terms of basic personality characteristics or motivations. The fact that the book has 1,032 foot-notes, many of them multiply documented, but scarcely a single genuine case study, correctly indicates the virtual absence of what usually is regarded as an integrative and dynamic point of view. Another result of the method is that phenomena of behavior which are descriptively separable, e.g. conditioning, trial and error learning, and insight, remain unrelated so far as the application of any kind of general psychological theory is concerned.

Another characteristic of the book is the citation of concrete examples of abnormal behavior to a degree not approached in any other book with which the reviewer is familiar. Many of these are in the form of anecdotes derived from the older

literature in the field. Thus the examples of illusions include quotations from Edmund Parrish (1888), T. H. Pear (1922), Janet (1895), Esquirol as reported by Maudsley (1880), Griesinger as reported by Ribot (1898), and Sully (1897). Many of the anecdotes could well be termed psychological curiosities. For example, on p. 381, in connection with the sub-topic of hallucinations, we find these quotations: "A carpenter was subject to visions of his dead child 'against his wish and will' whenever he concentrated his attention upon any point . . . Kahlbaum, quoted by G. Störring, *Mental Pathology in its Relation to Normal Psychology*, 1907." "In many instances, the apparition of a murdered person has haunted the murderer until he gave himself up to the police . . . Abercrombie noted this . . . *Inquiries*, 1833, 260." "The reader can only imagine the years of library research which must have been necessary to assemble this rather astonishing "wealth of material" (publisher's phrase).

To the reviewer it seems uncertain just what teachers of undergraduate abnormal psychology would find the book acceptable as a text. For an instructor to regard it as satisfactory, he must of course agree with the point of view and method of presentation already described. In particular he must be content with a treatment of the subject which remains almost entirely at the descriptive level and makes little or no attempt to explain or interpret abnormalities of behavior in terms, say, of an individual's motives and past history. He also must agree with Professor Taylor's selection of sources, which limits Fromm, Horney, and Adler to 6 or 7 references apiece, but has 32 such references to William James, 37 to Janet, and 46 to Morton Prince.

New York University

LELAND W. CRAFTS

The Psychiatrist and the Law. By WINFRED OVERHOLSER. New York, Harcourt, Brace and Company, 1953. Pp. x, 147.

The author, who is the superintendent of St. Elizabeth's Hospital, was the recipient of the Isaac Ray Award given by the American Psychiatric Association for his contribution to the improvement of the relations between Law and Psychiatry. This book consists of four lectures which he gave at Harvard University in fulfillment of the expectation that the recipient of the award would lecture at some university which has both a law and medical school.

In the first lecture there is a tribute to Isaac Ray for his contributions made over a century ago to the problem of medical jurisprudence, a brief account of the history of psychiatry as a medical specialty, and a general statement of the various symptoms and disorders with which psychiatry deals. Although he is catholic in his views, paying due regard to physiological disfunctions where they are known, the author nevertheless accepts the Freudian doctrine, even to the *id*, *ego*, and *superego*, admitting however that these terms have only a "schematic existence. They are convenient abstractions . . . and serve a useful clinical purpose" (13 f.).

The second lecture deals with the disinclination of the law to accept new points of view and "scientific data," especially the findings of psychiatrists. The author inveighs especially against the law's devotion to the dichotomy between sanity and complete insanity when in fact there are gradations between the two extremes for which the law should provide. Similarly, he deplors the tendency of legislative

bodies to react to public excitement with ill-considered laws especially "s xual psychopath legislation."

In the two remaining lectures he treats of the hospitalization of psychotics and of the psychiatrist as a witness. The commitment of a person to a public hospital is of course a legal responsibility. The author gives a general account of the wide variety of commitment-procedures in various jurisdictions. Some of these he finds to be exceedingly harmful to the patient who may be treated as a criminal, subjected to open-court proceedings, and forced to hear testimony against him by members of his family. On the other hand there are states that are adopting better methods. The problem of the psychiatrist as a witness is a difficult one. The author finds an antagonism between the courts and the psychiatric witness, which is due in part to the inability of the psychiatrist to present to the jury by means of charts and photographs the data upon which his opinions are based, and in part to the difference in approach—the lawyer bound by his rules, some of which are obsolete, the psychiatrist, unable to say whether the defendant knew right from wrong or could have appreciated the consequences of his act, knowing only that the defendant is mentally sick. Among other things the author pays his respects to the absurdity of the "hypothetical question."

Taken as a whole this is a good book and should be read by lawyers, judges, legislators, and the informed public. Its psychology, however, is typically psychiatric and there is no reference to the findings of those psychologists who are devoting their lives to the study of the basic psychology of the normal human adult. Otherwise the author could not say as he does that the Freudian psychology is "the generally prevailing psychologic doctrine today" (p. 19). Neither could he quote with approval the statement that "Every result of the use of the eyesight is as a matter of last analysis a deduction or judgment" (p. 57). Such statements are, however, entirely gratuitous and unnecessary to the general argument of the book. There is an index and a list of 118 references.

Cornell University

H. P. WELD

For a Science of Social Man. Edited by JOHN GILLIN. New York, Macmillan Company, 1954. Pp. vii, 289.

Social scientists have become self-conscious. They have long recognized that their common label is about all they have in common. Their enormous efforts and numerous publications both theoretical and empirical seem not to have influenced the course of human events nearly so much as the focused endeavors of the physicists, the professional soldiers, or the ward politicians elevated to national positions. Now there is a growing belief that if the social scientists could analyze their own thinking and establish rapprochements with each other, the net effect would carry us closer to an understanding of man in his own creation—society. This book is an effort at such self-analysis by a group of distinguished representatives from anthropology, sociology, and psychology.

The editor, Dr. Gillin, University of North Carolina anthropologist, provides an introductory and a concluding chapter. Each of the other six contributors deals with his own and with a neighboring 'discipline' in an attempt to highlight theoretical concepts, findings, or issues which in his judgment enhance the scope of the selected fields. Two weekend conferences, supplemented by correspondence,

were held by the authors as a means of bringing their thinking and consequently their writing into closer unity. What emerges from these deliberations is not a grand unified theory of social man—such was not their objective—but rather a group of thoughtful and provocative essays with more emphasis on theory than man. One wonders what intellect will ever grasp the totality of existing knowledge in this area, bringing direction and focus to our current scattered efforts.

Nearly all the contributors make mention, passing or otherwise, of certain concepts—culture, group-structure, rôle-behavior, personality, learning, communication. Is it possible that a greater integration might now be achieved if the same authors were asked to pool their thinking on these topics alone, each presentation organized around an agreed-upon outline? The authors, as one might expect, do not address themselves to each other. Their essays are like monologues, each man talking from the same theater-in-the-round but in a different direction and perhaps to a different segment of the audience. A rough check of the name-index indicates that their respective platforms infrequently overlap. No one person's works are cited by all contributors. To take a sample, Harry Elmer Barnes is quoted 13 times by two contributors, Freud 18 times by three contributors, Kroeber 30 times by three contributors, Malinowski 17 times by four contributors, Murdock 14 times by three contributors. Moreover, only one of the contributors refers to each of his fellow authors in this book. This survey suggests not only that the collaborators have a quite small common denominator of factual knowledge but that they are not overly influenced or acquainted with each other's publications. This observation does not detract from the high level of scholarship evidenced, but is offered instead as evidence that the higher integration sought for can be attained only after much diligent, long-term distillation of ideas and considerably more opportunity for face to face discussion than could be afforded to by two weekend conferences.

This problem has not escaped the authors. In one fashion or another nearly all have recognized the barriers to communication posed by sheer limitations of time, problems of semantics, departmentalization of faculties, and channels of publications. Some specialists, furthermore, are temperamentally unsuited to interdisciplinary collaboration. Competition has, after all, a high valence in the American scene, among scholars as well as students and business men. Newcomb, among others, points out that the integration of concepts does not, however, necessarily imply the loss of any specialty's "sovereignty." "The present requirement for a unified science of man is not a fusion of the different levels of inquiry, but recognition and understanding of what is identical and what is distinctive. From such recognition and understanding, theoretical integration may yet emerge."

This book will be more respected than read. The writing of it has probably taught its contributors more than it will teach its readers. It is an important and necessary piece of scholarship which anyone concerned with the integration of social sciences should examine carefully for the clarification of his own views and for the interdisciplinary research which it suggests. It represents a necessary step for a science of social man.

Institute for Research in Human Relations

F. K. BERRIEN

Child Training and Personality. By JOHN W. M. WHITING and IRVIN L. CHILD. New Haven, Yale University Press, 1953. Pp. vi, 353.

This book is an attempt to apply Hullian theory to the analysis and testing of

in relationships between child-training practices and adult personality.⁹ The relationships with which the book is concerned were suggested by psychoanalytic theory, but are cast into terms of behavior theory and tested by reference to ethnographic reports on 75 primitive societies from Yale's Human Relations Area File. The translation of cultural data into terms developed in a theory of individual behavior is effected by introducing the concept of custom as the unit of culture, and by defining custom as "a characteristic habit of a typical member of a cultural category of persons" (p. 22).

The customs chosen for analysis concern child-training practices and beliefs about causes and cures of illness. Relevant data were collected from the Area Files and presented for rating to three graduate students. Five systems of behavior were selected for consideration: oral, anal, sexual, dependent, and aggressive. Ratings and rankings were made on the degree of initial satisfaction permitted by the societies for each system of behavior, and on the degree of anxiety engendered by attempts to socialize each system. Two additional students rated each society on the degree to which it attributed the cause and cure of illness to each of the five systems of behavior.

The hypotheses concerning the relationships between these customs were derived from the Freudian discussion of fixation. A distinction is made between positive fixation, which is assumed to result from a high degree of initial satisfaction, and negative fixation, resulting from a high degree of socialization-anxiety. With regard to illness, it is hypothesized that the causes will be attributed by primitive societies to behavior which is negatively fixated, and cures will be sought for behavior which is positively fixated. In the case of oral behavior, strong support for the hypothesis of negative fixation was found, for the ratings of oral socialization-anxiety were positively related to the ratings of oral explanations of illness. Similar confirmation was obtained for aggression, but not for the anal, genital, and dependent systems. The positive-fixation hypothesis was not supported, for ratings of initial satisfaction were not related to the incidence of oral, anal, or similar therapeutic practices. Some evidence was found for positive fixation, however, when behavior appearing after early childhood was considered. Finally, the origins of guilt and fear of others are discussed. The societies' conceptions of illness are used to show that the tendency to attribute the cause of illness to the patient arises from identification with punishing parents, and that the tendency to attribute the cause of illness to other persons is associated with early anxiety about aggression. It is of interest to note that ratings of the child-rearing practices of middle-class American families showed these families to be more severe and less indulgent in all respects of oral, anal, and genital training than the parents in primitive societies. They were nearer the median of the primitive groups with regard to dependence and aggression.

The task which the authors have undertaken is obviously of great magnitude and complexity, and the paucity of positive results is not surprising. Considering the difficulties with which they had to deal, it is surprising that any positive results were achieved at all. The authors discuss their study with great caution; indeed the concepts and findings are so amplified and qualified that their meaning frequently becomes obscured. This tendency, in addition to the authors' difficult literary style, makes the book hard to read. Those who do read it, however, will find it a

stimulating exercise in the extension of concepts and hypotheses from behavior theory into a new and interesting domain.

University of Texas

HAROLD W. STEVENSON

Motivation and Morale in Industry. By M. S. VITELES. New York, Norton, 1953. Pp. xvi, 510.

Industry continually needs to increase production, to promote employee satisfaction and adjustment at work, and to curtail industrial strife. "These objectives cannot be fully achieved . . . unless accurate information is available as to the effectiveness of the various appeals or incentives which can be used to arouse the cooperation of individual workers. This involves an exact determination of basic motives in work, in the sense of evaluating the nature and strength of various wants and needs. . . ." To this end Dr. Viteles brings to bear a wealth of experimental studies and attitude surveys. Although the book is oriented toward industry, the academician should not be misled by its title. Half of the 1500-or-so references are to academic sources, and many instructors will find the book useful in the teaching of motivation.

Dr. Viteles frames his problem in a section on "Mobilizing the Will-to-Work," noting that attention must be focused on the individual, and describing wage-incentive systems and their relation to the worker. The second section, "The Background of Motivational Theory," summarizes classical theories of motivation, the contributions of social psychology, and the work of Lewin and his followers. The section on "Experimental Studies" reviews the relations between productivity and such variables as job-preference, size of group, supervision, participation in decision making, social functions of the working group, and teamwork. The largest section, "Employee-Attitude Surveys," begins with a discussion of the methodology of attitude surveys, and then considers the results of surveys of attitudes on such questions as unions, pay envelope, status, and various plant features. Finally, there are the "Guideposts for Management." The problem of motivation in industry," says Dr. Viteles, "is one of energy mobilization. The problem of morale is that of mobilizing the sentiments of all involved in achieving the objectives of industry within the broader framework of society as whole." Here he directs management's attention toward the crucial variables involved in mobilizing the energy and sentiments "of all involved."

This book will probably find its way to the bookshelves of all major industrial executives and of all industrial psychologists. Those experimental, social, and clinical psychologists who take the time to read it will find interesting new ways of considering their own motivational theories and problems.

Knox College

ROBERT S. HARPER

The Measurement of Hearing. By I. J. HIRSH. New York, McGraw-Hill Book Company, 1953. Pp. 364.

The author meant this book to be used "as a reference by those engaged in measuring and treating disordered hearing and also as a textbook by those who are training for such clinical work." He also intended it "to be used by those students of psychology, physiology, and physics who have an interest in hearing."

The first three chapters, devoted to "Sensation and Measurement," "The Auditory

Stimulus: Sound and Electricity," and "Electroacoustic Systems," provide a review of these subjects which should serve as a convenient reference for students already familiar with the material. They would prove rather difficult for the beginner. Chapter 4 is about pure-tone audiometry. It describes various parameters which affect threshold measurements, presents data from several surveys, and discusses a number of audiometric techniques. Chapter 5, concerned with the intelligibility of speech, discusses speech sounds, types of speech, articulation testing, speech audiometry, and the index of social adequacy. Combined with Appendix B, which contains several spondee and PB lists taken from the Psychoacoustic Laboratory series and the CID series, the chapter should prove a valuable and convenient source of information on speech and speech-testing.

The next four chapters contain a wealth of up-to-date research material on masking, fatigue, differential sensitivity to pitch and loudness, loudness recruitment, binaural phenomena, and bone conduction. Since much of the research was done by the author himself, and since he was closely associated with many of the workers responsible for the rest, he writes with a genuine understanding of the material and of its significance for the clinician and audiologist.

The next chapter deals with the controversial topic of conditioning in audiometry, and raises the provocative question "where does the auditory system end?" The final chapter is entitled "Clinical Audiometry" and includes a discussion of the aims and methods of clinical audiometry, of the meaningfulness of "over-all estimates of hearing," and of various therapeutic measures.

There are two appendices, the one already mentioned and another containing the ASA standards for audiometers of various types. There is also a glossary of technical terms, a bibliography, and a name- and subject-index. Each chapter contains a short annotated list of references for further study.

The book should prove a rich source of information for anyone interested in the measurement of hearing and hearing-loss. The audiologist will not think his task an easy one after he has read the book, but he will be able to avoid many of the pitfalls into which he might otherwise have fallen. The student of psychology will discover, among other things, that he should have learned more physics.

University of Texas

L. A. JEFFRESS

Social Psychology. By OTTO KLINEBERG. Revised Edition. New York, Henry Holt and Company, 1954. Pp. xiii, 578.

With this revision Klineberg has brought up to date a work that deservedly ranks as one of the most useful textbooks of social psychology. Clearly and interestingly written, well documented, intelligently selective in its reports of research, consistent in its theoretical position without being too heavily theoretical, Klineberg's book may well stand as a model of textbook writing. Several generations of students have been nurtured on the first edition. One predicts without hesitation that the new edition, painstakingly revised and expanded to encompass new developments, will be received with equal enthusiasm.

Klineberg stands for a psychology that is not culture-bound. In every field of psychological inquiry, from "national character" down to the elementary processes of sense-perception, he insists that we can learn from the study of behavior in varying cultural contexts. There is possibly no psychologist who has done so much,

through his teaching, his research, and his extra-curricular activities, to bridge the gap between psychology and ethnology, and to facilitate international understanding among psychologists. Klineberg gives us a picture of man as a highly adaptable creature, his behavior initiated perhaps by certain dependable motives, but its expression regulated in all its essentials by his culture. The psychological framework is, broadly speaking, that of Woodworth; the cultural emphasis reveals the influence and the idealism of Boas. Some readers may find the relativism a little too 'easy,' but none can deny that Klineberg's approach both fosters tolerance and stimulates the urge to find out more about the behavior of people.

The revised edition contains two new chapters that are particularly noteworthy: one on "The study of national characteristics," in which Klineberg capably reviews the research and points up the issues in this perennially frustrating field; the other on "Psychology and international relations," which, in spite of its measured tone, is a ringing challenge to psychology to play its part in the illumination of world problems. Since World War II, psychologists have been beginning to appreciate their rôle as scientists in the solution of the larger problems of mankind. Klineberg's *Social Psychology*, whether studied as a text or simply read as an interesting analysis of human behavior, will add to the challenge.

Cornell University

R. B. MACLEOD

Einführung in die Sozialpsychologie. By PETER R. HOFSTÄTTER. Vienna, Humboldt-Verlag, 1954. Pp. ii, 535.

This introduction to the general field of social psychology is the first one to appear in German which contains an adequate summary of experimental investigations carried out in the Anglo-Saxon countries. From the point of view of content it is not very dissimilar from the major textbooks used at American universities; it is no surprise to find that it is based on courses of lectures given by the author at M.I.T. and the Catholic University of America.

While there is thus little that is novel in the book to an American reader, most of the material would be quite unfamiliar to the audience for which it is intended, and teachers in German speaking countries will undoubtedly be glad to have this summary and integration of experimental work to recommend to their students. Indeed, it is likely that a good deal of the contents will be news to the teachers as well!

Certain features of the book call for comment. In the first place, it is extremely well written. The reviewer knows of no comparable volume in which clarity of exposition is so happily married to felicity of expression, and in which so much factual information is conveyed in a context of humanistic background. In the second place, this appears to be the only textbook of social psychology to grasp firmly the nettle of factor analytic and other complex statistical methods and to introduce them as an essential part of the presentation. The necessity of doing this has always impressed the reviewer in his own teaching activities, and it is a pleasure to find this difficult task done so well. A third feature which calls for comment is the author's wide reading, which enables him to treat even well-worn topics with an air of novelty and with amusing and unusual examples and references.

Altogether, this is an excellently constructed, well written textbook, which is at

least the equal of any available in English at the present time. Its defect, a lack of basic theoretical integration, can probably be laid at the door of the subject more fittingly than at that of the author. An English translation of the book might make it into a formidable rival of more familiar textbooks.

Maudsley Hospital

H. J. EYSENCK

Psychosomatic Case Book. By ROY R. GRINKER and FRED P. ROBBINS. New York, The Blakiston Company, 1954. Pp. xii, 346.

The description and analysis of 79 psychosomatic cases take up the bulk of this book. Intended primarily for students of medicine and related disciplines, the case reports are grouped into eight classes or special syndromes such as cardiovascular, respiratory, and so forth. The cases vary in length from half-a-page to four pages, and in quality from fragmentary notes to carefully structured histories.

The case reports are intended to give the reader some familiarity with the authors' "field concept of psychosomatic medicine," the exposition of which (in Chapter III) constitutes what is by far the best material in the book. The authors provide argument together with general data and case material leading to the rejection of the popular simplicistic theories which hold that the psychosomatic field is an interrelationship "between two foci or systems of variables such as specific emotions and specific organ functions, specific life stress and a bodily reaction. . . ." Instead, this book emphasizes the notion of multiple causation. In eschewing the use of specific stereotypes or symbolic vegetative conditions in favor of a "transactional" approach, the authors are to be commended. Nevertheless, the symbolism of bodily organs creeps silently into some of the case discussions.

Although one should not expect detailed exposition of theory in a case book, the reviewer believes that the presentation would have been considerably more valuable if the transactional and field-theoretical principles had been systematically employed in the case analyses.

University of California

THEODORE R. SARBIN

The Measurement of Social Competence: A Manual for the Vineland Social Maturity Scale. By EDGAR A. DOLL. Philadelphia, Educational Publishers, 1953. Pp. 664.

The Vineland Scale, first published in 1936, represents a unique attempt to measure social maturity in terms of data elicited from an informant well acquainted with the everyday behavior of the person to be evaluated. The volume at hand—crammed with a mass of detailed material on the rationale, construction, validation, administration, scoring, and use of the scale—replaces the brief and inadequate manual which was, up to the present time, the only available guide. The impression is gained of a careful, competent piece of work which makes a significant contribution to the evaluation of the physically handicapped, the illiterate, and the mentally retarded. It is recommended to those counsellors who routinely base their judgments on scores derived from intelligence tests alone.

University of Texas

IRA ISCOE

INDEX

By H. W. STEVENSON, University of Texas

(The names of authors of original articles are printed in CAPITALS; of authors of books reviewed, in roman; and of reviewers, in *italics*.)

- Abbott, B. C. 383
ADAMS, J. A. 268
ARAMS, P. A. 338,
..... 513, 612, 708
AIKEN, E. G. 26
Allport, G. W. 742
Anderson, I. H. 197
ANDREWS, T. G. 453, 704
Asch, S. 371
Ashby, W. R. 391
ATKINSON, R. L. 521
AUSTIN, P. 338,
..... 513, 612, 708
BACHEM, A. 360
BARTLEY, S. H. 500
BEACH, F. A. 176
BECK, L. H. 152
Bellak, L. 393
BERGER, C. 509
Berrien, F. K. 739, 748
BILLINGSLEY, B. A. 335
BINDRA, D. 554
BITTERMAN, M. E. 129
..... 133, 205, 335, 427,
..... 525, 717
Bitterman, M. E. 394, 743
Blake, R. R. 393
BORING, E. G. 369,
..... 370, 573
Boring, E. G. 192,
..... 204, 383
BOUSFIELD, W. A. 111
BRIDGES, C. C. 525
BRIGGS, G. E. 475
BROGDEN, W. J. 475
Brown, D. R. 203
BROWN, K. T. 533
BROWN, R. H. 104
BROŽEK, J. 590
Brunswik, E. 386
BURNHAM, R. W. 492
CARMICHAEL, L. 732
Carroll, J. B. 568
Carroll, J. B. 744
Child, I. L. 749
Christie, R. 742
CHRISTMAN, R. J. 484
COFER, C. N. 453
COHEN, B. H. 111
Cohen, L. H. 390
COHN, V. H. 124
COLGAN, C. M. 711
Conant, J. B. 192
CONKLIN, J. E. 104, 363
Crafts, L. W. 388, 746
CRAWFORD, F. T. 717
CRONBACH, L. J. 368
CRUMBAUGH, J. C. 647
DALLENBACH, K. M. 178,
..... 571, 734
DAVIS, R. C. 343
DAY, R. H. 659
DEATHERAGE, B. H. 259
Dearborn, W. F. 197
DEESE, J. 299
DINGMAN, H. F. 395
Doll, E. A. 754
DURYEA, R. A. 427
EDWARDS, W. 56
..... 68, 164, 441
ELAM, C. B. 133
Eysenck, H. J. 753
FEDDERSEN, W. E. 335
Fenlason, A. F. 394
FERGUSON, G. A. 178
Ferguson, L. W. 195
FERNBERGER, S. W. 288
Floyd, W. F. 567
Gardner, M. 204
Geldard, F. A. 383
GHENT, L. 220
Gillin, J. 748
Gilmer, B. v. H. 392
Gilmer, B. v. H. 202
Gray, J. S. 392
GREEN, E. J. 141
Grinker, R. R. 754
GROSS, N. B. 124
GRUBER, H. E. 411
GUILFORD, J. P. 395, 691
Guilford, J. P. 195
Guthrie, E. R. 394
HAGGARD, E. A. 357
HAHN, E. L. 500
HALL, J. F. 138
HANAWALT, N. G. 170
HARDMAN, G. W. 299
Harper, R. S. 202,
..... 570, 751
Harris, D. B. 198
HARRISON, J. M. 353
HARTMAN, E. B. 1
HELSON, H. 321, 327, 677
Helson, H. 189
Hilgard, E. R. 393, 561
Hirsh, I. J. 751
HOCHBERG, J. E. 205
Hochberg, J. E. 386
HOFFMAN, E. L. 143
Hofstätter, P. R. 753
Holtzman, W. H. 566
Horava, A. 394
HOWES, D. 241
Hull, C. L. 375
Iscue, I. 754
JAFFE, R. 668
Jeffress, L. A. 751
JENKINS, J. J. 700
JOHNSON, H. M. 723
JONES, F. N. 147
JONES, M. H. 696
KARLIN, L. 358
Karn, H. W. 392
Kecksemeti, P. 744
KENNEDY, J. L. 367
Kilpatrick, F. F. 379
KING-ELLISON, P. 700
Kleemeier, R. W. 200
Klineberg, O. 752
KRAUSKOPF, J. 205,
..... 278, 427, 684
Krauskopf, J. 204, 391
Krech, D. 561
Kubie, L. S. 393
LACEY, O. L. 369
Lamb, G. 388
LALÉ, G. 565

- LAW, O. T. 714
 Lawrence, M. 735
 LeCron, L. M. 199
 Leeper, R. 375
 LEIBOWITZ, H. 530
 LOCKE, N. 341
 Lowy, K. 735
 • LYBRAND, W. A. 704

 MACCASLIN, E. F. 308
 MacLeod, R. B. 371, 752
 MAHNEKE, A. 509
 Maier, N. R. F. 202
 Marcuse, F. L. 199, 389
 Mathews, R. W. 182
 MAYER, C. E. 353
 MAYES, G. L. 717
 MCCONNELL, J. V. 129
 McKinney, F. 392
 Menninger, K. A. 391
 MEYER, M. F. 39, 538
 MICHEL, W. C. 321,
 327, 677
 Moore, F. J. 393
 MOWRER, O. H. 15, 26
 MULHOLLAND, T. B. 546
 MULL, H. K. 341
 Murray, E. 182
 Myers, C. A. 200

 NAFE, J. P. 179
 NEWHALL, S. M. 96
 Nuttin, J. 388

 Ogden, R. M. 194, 568
 ORD, N. 341
 OSGOOD, C. E. 241
 Osgood, C. E. 555

 Overholser, W. 747

 PAGE, H. A. 154
 Pastore, N. 379
 Pickford, R. W. 189
 Piéron, H. 383
 Pigors, P. 200
 Pirenne, M. H. 383
 Pollak, O. 198
 PORTER, P. B. 550
 POSTMAN, L. 119, 612
 Powell, J. N. 190
 PRATT, C. C. 551
 PRENTICE, W. C. H. 315,
 727
 Prentice, W. C. H. 555
 Pumpian-Mindlin, E. 393

 REYNOLDS, B. 268
 Rao, C. R. 566
 ROBBINS, F. P. 754
 ROBINSON, E. J. 464
 ROGERS, L. S. 366
 ROHRER, J. H. 143
 Ross, S. 704
 Ryan, T. A. 739

 SALTZMAN, I. J. 521
 SANFORD, F. H. 177, 730
 Sarbin, T. R. 754
 SEDGEWICK, C. H. W. 111
 Selye, H. 393, 394
 SEMMES, J. 220
 SHAPIRO, S. 453
 SIDONS, G. F. 343
 Slotkin, J. S. 203
 SMITH, K. 174
 Smith, P. C. 739

 Smith, P. C. 394
 SOLOMON, L. N. 15
 Spinley, B. M. 570
 STOUT, G. L. 343
 Strecker, E. A. 392
 STURGEON, A. 321

 TAYLOR, H. E. 590
 Taylor, W. S. 746
 TEUBER, H. L. 220
 THOMAS, G. J. 632
 THOMPSON, R. F. 475
 THOMPSON, W. R. 357
 Tiffin, J. 202
 Tinker, M. H. 197
 TUMA, A. H. 119
 TYLER, L. E. 553

 Vanderplas, J. M. 567
 Vernon, M. D. 204
 Vetter, G. B. 190
 Vinacke, W. E. 194
 VINCENT, D. F. 155
 Viteles, M. S. 751

 WALLACH, H. 338, 513
 Walls, G. L. 182
 WARREN, J. M. 517, 720
 WEINSTEIN, S. 220
 Weitzenhoffer, A. M. 389
 Weld, H. P. 390, 747
 Welford, A. T. 567
 WERTHEIMER, M. 543
 Wever, E. G. 735
 Whiting, J. W. M. 749
 Worchel, P. 391
 WYCKOFF, L. B. 154

SUBJECTS

(References in *italic* figures are to reviews.)

- Abnormal psychology, 393, 746 f.
 Absolute judgments, effect of anchors on, 12 f.
 Absolute method, *see* Method of single stimuli.
 Absolute pitch, 1-14.
 Acoustics, 735 ff.
 Acknowledgment of Hunter's photograph, 734.
 Adaptation, level, 321 ff.; 327 ff.; syndrome, 394.
 Adjustment, method in psychophysics, 546-550.
 Affective value, auditory stimuli, 691-695.
 After-effects, brain lesions, 668-678; figural, 259-267, 684-690; 'attraction' in, 174-176, errors in measuring, 543-546, stroboscopic movement and, 259-267; kinesthetic, 668-676.
 Aim, accuracy of, 659-667.
 American Association for the Advancement of Science, *see* Meetings.
 American Philosophical Society, *see* Meetings.
 American Psychological Association, *see* Meetings.
 Anchoring effects, judgments of pitch and, 12; operant responses of rats, 141 f.
 Animal, monkey, discrimination, 517-520; discriminative learning, 720-722; rat, conditioned fear, 15-25, 26-38; discrimination, 133-137, 308-314, 335-337; discriminative learning, effects of, 717 ff.; hoarding, 124-128; operant responding, 141 f.; stripe preferences of, 714-716; successive discrimination, 129-132.
 Apparatus, absolute pitch, 1-14; arterial pressure, 351-352; autokinetic movement, 525-529; autonomic activity, 343-352; CFF, 634; gambling, pinball, 57, 69; GSR, 343-347, 357 f.; light, movement of, 143-146; 'Maxwellian view', 530 ff.; movement, visual, 105; non-gambling, 81; olfactometer, 147-151; plethysmograph, electrical, 348-351; pulse, 351-352; shock grid, 154; size-constancy, 464 f.; stimulator, square wave, 353-356; timer, electronic, 358 f.; visual, quantal fluctuations, 152 f.
 Apparent movement, figural after-effects, 174 ff., 259-267; measurement of, 525-529.
 Apparent orientation, object, 500-508.
 Apparent size, brightness, 464-474; figural after-effects, 543-546; fluctuations in, 533-538.
 Audition, auditory stimuli, affective value of, 691-695; fatigue, 538-543; pitch shifts, 484-491; figural after-effects in, 278-287; hydraulic theory, 39-55; identification of pitch, 1-15; measurement, 751 f.; physiological acoustics, 735-739.
 Autokinetic movement, 525-529.
 Beats, multiple frequency, 39-55.
 Behavior systems, 375-379; Hull's goal, 375 f.; value of Hull's principles, 378 f.; transitional point of view, 379-383.
 Bets, probability-preferences, 56-67, 68-95.
 Binocular, rivalry, 513-516; subjective colors, 492-499.
 Biography, Walter Samuel Hunter, 732-734; Harry Miles Johnson, 179-181.
 Brain, design for, 391; functions, 220-240; injury, 220-240.
 Brightness, effect of, on reversible perspective, 341-342.
 Business psychology, 200 f., 392 f., *see* Industrial psychology.
 Canadian Psychological Association, *see* Meetings.
 Cerebral lesions, kinesthetic after-effects following, 668-676.
 Chemical sense, the 'common' 696-699.
 Child, psychotherapy, 198 f.; training, personality and, 749 ff.
 Children, rote learning in, 453-463.
 Clinical psychology, case-reports, 391 f.; research, 743 f.
 Cochlear mechanics, hydraulic theory of, 39-55.
 Color, binocular, 492-499; binocular rivalry of, 513-516; blindness, means of studying, 182 ff.; vision, foveal, 182 ff.
 Conceptual framework of psychology, 386 ff.
 Conditioning, fear, 12-25, 26-28; operant, 141 f.
 Contiguity vs. drive-reduction, 15-25, 26-38.

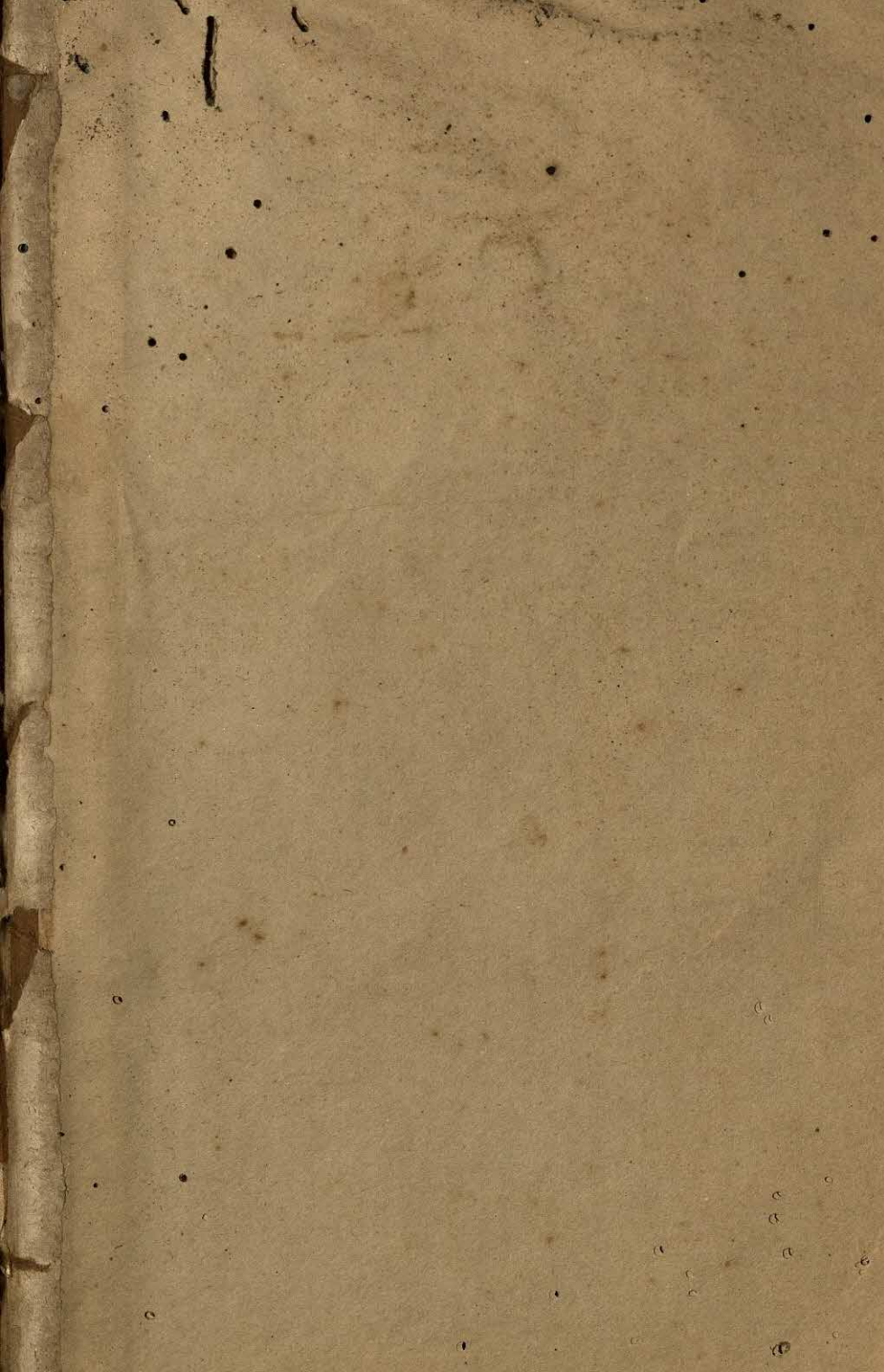
- Control, experimental, 573-590.
 Cross-sensory judgments, 13.
 Critical flicker frequency, age, intelligence, and, 711 ff.; fatigue and, 509-512; interocular stimulus differences and, 632-646.
- Deficiency, vitamin-B, 124-128.
 Diffusion model, visual form threshold, 205-219.
- Discrimination, after training on white-black, 714-716; nondifferential reinforcement, effects on, 335-337; nondifferential reinforcement and, 133-137, 717-719; rate of, monkey, 517-520; reversed, 720-722; successive, 129-132; successive and simultaneous, 129-132, 308-314.
- Drive-reduction, proximity and abruptness, 15-25; temporal variations, 26-38.
- Dynamic psychology, 746 f.
- Eastern Psychological Association, *see* Meetings.
- Editorial announcement, 571 f.
- Errata, 178.
- Experimental psychology, controls, nature and history, 573-589; method and theory of, 550-560.
- Factor analysis, 155-163.
- Fatigue, auditory, 538-543; perceptual organization and, 704-707; symposium on, 567 f.; visual, 509-512.
- Fear, conditioned, 15-25, 26-38.
- Fechner's, colors, 492-499; law, 406 f.; VEG scale and, 677-683.
- Figural after-effects, 'attraction' in, 174-176; auditory, 278-287; duration of test-period and, 684-690; errors in measuring, 543-546; kinesthetic, 668-676; stroboscopic movement and, 259-267.
- Fitness, motor functions and, 590-611.
- Fluctuation, reversible perspective, 341-342, 708-710; size, apparent, 533-538.
- Forgetting, absolute pitch, 10.
- Form, visual, threshold for, 205-219.
- Frequency, auditory, affective value, 691-695; critical flicker, *see* Critical flicker frequency.
- Gambling, probability preferences in, 56-67, 68-95; uncertainties, methods for computing, 164 ff.; variance preferences in, 441-452; vs. non-gambling risks, 81-94; with real money, 75; with worthless chips, 74.
- Generalization, rote serial learning, 453-463.
- Habits, visual discrimination (monkey), 517-520.
- Hearing, hydraulic theory, 39-55; measurement of, 751 f.; physiological acoustics, 735-739.
- History, editorial, this JOURNAL, 571-572; experimental control, 573-589.
- Hoarding, vitamin deficiency and, of rats, 124-128.
- Hunter, W. S., *see* Biograph; photograph, facing p. 571.
- Human relations, principles of, 202 f.
- Hypnosis, experimental, 199 f., 389 f.
- Hypotheses, verification of, 723-727.
- Illusion, apparent size, 533-538; puzzle picture, 550 f.
- Images, recurrent, 170-174, 360 f.
- Incidental learning, intentional learning and, 521-524.
- Industrial psychology, 200 f., 202 f., 392 f., 394, 739 ff., 751.
- Inhibition, retroactive, 299-307.
- Inhibiting potential, transfer and generalization of, 453-463.
- Intensity, auditory, affective value of, 691-695.
- International Congress of Psychology, *see* Meetings.
- Introduction to psychology, 561-565.
- Johnson, Harry Miles, *see* Biography.
- Kinesthesia, after-effects of cerebral lesions, 668-676.
- Korte's laws, stroboscopic movement, 259-267.
- Language, 568 f.; associative probabilities in, 241-258; relation to visual recognition, 315-320.
- Law, of effect, experimental analysis of, 612-631; psychiatry and the, 390 f., 747 f.; *see* Legal.
- Learning, 394, 557 f., 743 f.; conditioned fear, 15-25, 26-38; discriminative, 129-132, 133-137, 308-314, 335-337; discriminative (monkey), 517-520, 720-722; (rat), 717-719; effectiveness of rewards and punishment, 619-624; incidental vs. intentional, 521-524; latent, 119-123; law of effect, 612-631; massed vs. distributed practice, 268-277; operant, 141 f.; psychomotor, 268-277; rote serial, 453-463; rote verbal, 299-307; theory, 743 f.; tracking, 475-483; verbal associates, 111-118; word-frequency and, 138-140.
- Legal psychology, 390 f., 747 f.
- Localization, visual traces, 338-340.
- Logical fallacies, 723-727.

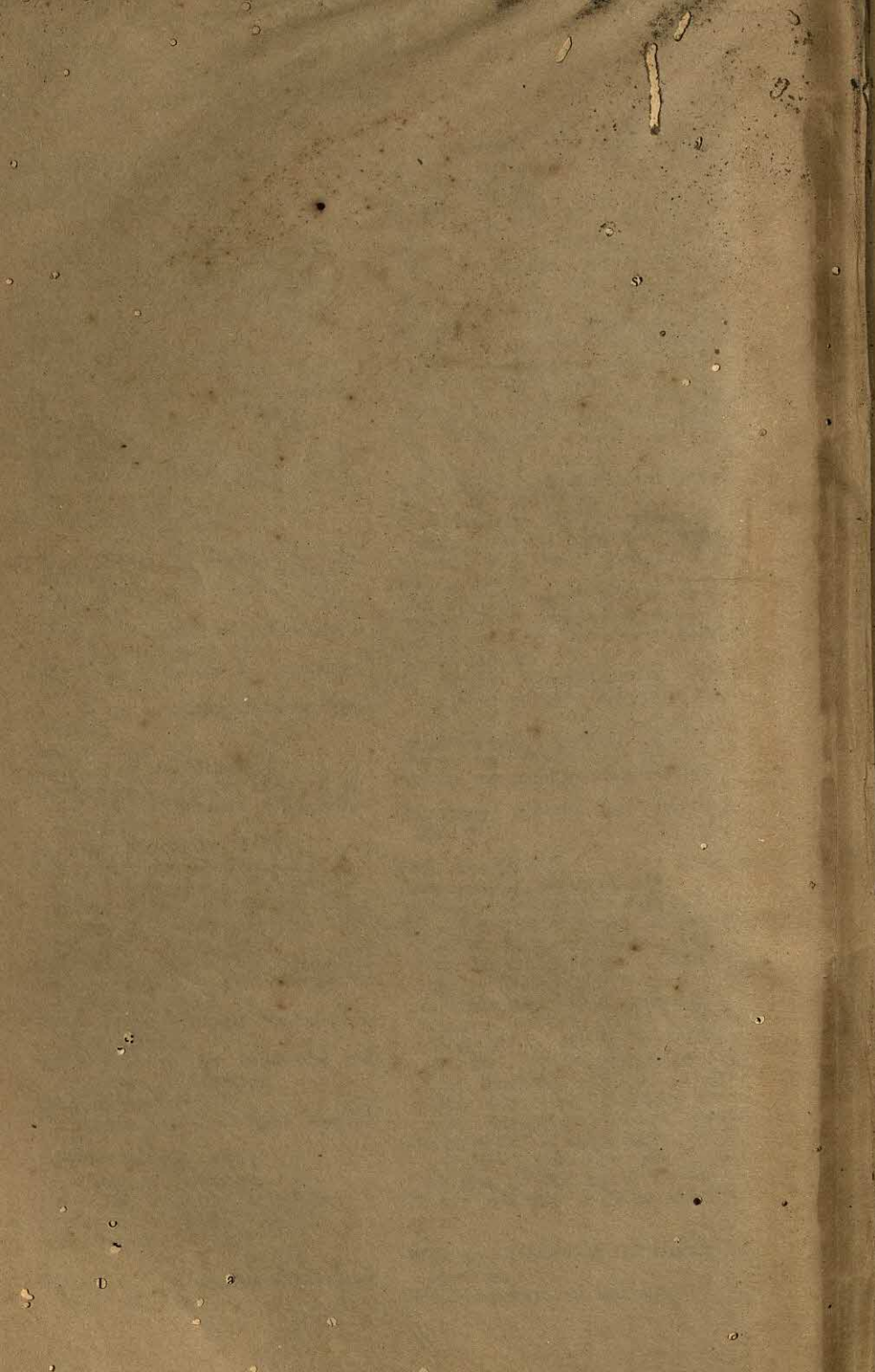
- Madness and the law, 390 *f.*
 Man, modern, science and, 192 *ff.*
 'Maxwellian view,' 530-532.
 Maxwell's spot, 182 *f.*, 185.
 Meaning, philosophy of, 744 *ff.*
 Measurement, personality, 195 *f.*
 Meetings, American Association for the Advancement of Science, Sec. I, 176 *f.*; American Philosophical Society, 370; American Psychological Association, 177 *f.*, 730 *f.*; Canadian Psychological Association, 178, 554; Eastern Psychological Association 365 *f.*; International Congress of Psychology, 551 *ff.*; Midwestern Psychological Association, 368 *f.*; National Academy of Sciences, 369 *f.*; Rocky Mountain Branch of American Psychological Association, 366 *f.*; Society of Experimental Psychologists, 367 *f.*; Southern Society for Philosophy and Psychology, 369; Western Psychological Association, 553 *f.*
 Memory, musical, 360-363; verbally labelled figures, 315-320; visually perceived form, 647-658.
 Method, of adjustment, 546-550; of agreement, 574; of concomitant variations, 575; of difference, 574; of paired comparisons, 96-103; of single stimuli, 96-103, effect of anchors on, 12 *f.*; of successive reproduction, 647; statistical, 566 *f.*
 Midwestern Psychological Association, *see* Meetings.
 Model, brain, 391; visual form, diffusion, 216 *f.*; Osgood-Heyer's, 218.
 Monkey, *see* Animal.
 Motor functions, accuracy of aim, 659-667; tests of, human fitness and, 590-611.
 Movement, apparent, 174 *ff.*, 259-267, 525-529; autokinetic, 525-529; visual thresholds of, 104-110.
 Murder and the law, 390 *f.*
 Musical memory-mechanisms, 360-363.
 National Academy of Sciences, *see* Meetings.
 Obituary, *see* Biography.
 Olfactometer, 147-151.
 Opinion, public, 190 *ff.*
 Operational point of view, 193, 195.
 Orientation in darkness, 500-508.
 Practice, effects of, 594; influence on absolute pitch, 1-14.
 Perceived form, temporal change in, 647-658.
 Perceived size, photometric brightness and 464-474; relation to perceived distance, 411-426.
 Perception, auditory pitch, 404-491; binocular subjective color, 492-499; critical flicker frequency, 632-646, 711-713; following fatigue, 704-707; judgment of size, 464-474; object orientation, 500-508; pain, 696-699; perceived size 411-426; perceptual organization, 708 *ff.*, processes, 556 *f.*; reversible perspectives, 341 *f.*; tactual, 220-240; threshold for words, 700-703; transactional approach, 779 *ff.*; verbally labelled figures, 315-320; visual, 204, form, 205-219, 647-658.
 Personality, child training and, 749 *ff.*; development, 203 *f.*; measurement of, 195 *f.*; psychoanalysis and, 195 *f.*; subculture and, 570; theory, 743 *f.*; types, 570.
 Personnel, administration, 300 *f.*
 Perspective, reversible, 341 *f.*, 708-710.
 Phenomenal object, 500-508.
 Photometric brightness, relation to judgments of size, 464-474.
 Pitch, absolute, identification of, 1-14; shifts in, 484-491.
 Pitch-distance, influence on absolute pitch, 3-14.
 Portrait, W. S. Hunter, facing p. 571.
 Prejudice, nature of, 742 *f.*
 Probability-preferences in gambling, 56-67, in non-gambling risks, 81-94; reliability of, 68-95; with real money, 75; with worthless chips, 74.
 Propaganda, 190 *ff.*
 Psychiatry, basic, 342; case study manual, 391 *f.*; law and, 390 *f.*, 747 *f.*
 Psychoanalysis, personality and, 388 *f.*, science and, 393.
 Psychologists, American, prestige and impact, 288-298.
 Psychophysics, adjustment method, 546-550; comparative rating scales and, 321-326; judgments, 363-365; methods of single stimuli and paired comparisons, 96-103; ratio judgment-methods, 395-410; time-order effects, 327-334; VEG scale, 677-683.
 Psychosis, 313.
 Psychosomatic case book, 754.
 Psychotherapy, for children, 198 *f.*
 Public opinion, 190 *ff.*
 Pursuit-movements, 268-277, 475-483.
 Puzzle-picture, 550 *f.*
 Rat, *see* Animal.
 Rating scales, comparative, 321-326.

- Ratio, frequency of auditory beats, 40-55; judgment-methods, 395-410.
- Reading, teaching of, 197 f.
- Recall, verbal associates, 111-118.
- Recognition, visual, 315-320; function of word-frequency, 700-703; visual traces, localization of, 338-340.
- Recurrent images, 170-174, 360 ff.
- Reinforcement, nondifferential, 133-137, 335-337, 717 ff.
- Response-sequence, 363-365.
- Retroactive inhibition, 299-307.
- Reversible perspective, effect of brightness on, 341-342; effect of past experience on, 708-710.
- Rocky Mountain Branch of American Psychological Association, *see* Meetings.
- Rotary pursuit test, 268-277.
- Science, modern, 192 ff.; pseudo-, 204; psychoanalysis, 393; social man, 748 f.; theory of, 192 ff.; uses of, 204.
- Sensation, 'common' chemical, 696-699; functions, processes, and mechanisms, 383-386, 556; human, 385 f.
- Serial learning, transfer and generalization in, 453-463.
- Set, rôle in successive discrimination, 129-132.
- Shock, electrical, 154.
- Single stimuli, *see* Method of.
- Size, constancy, 411-426; effect of, on aim, 659-667; influence of brightness on, 464-474.
- Smell, an olfactometer, 147-151.
- Social psychology, 371-375, 752 f., 753 f., 754; deprived and privileged, 570; murder, madness, and the law, 390 f.; prestige of various psychologists, 288-298; public opinion, 190 ff.
- Social science, essays, 748 f.; prejudice, nature of, 742-743; psychotherapy for children and, 198 f.
- Society of Experimental Psychologists, *see* Meetings.
- Southern Society for Philosophy and Psychology, *see* Meetings.
- Space, auditory, 278-287.
- Spearman's formulae in factor analysis, 155.
- Statistical methods, 566 f.
- Stimulator, *see* Apparatus.
- Stress, 394; sensitivity to, 599-610.
- Stroboscopic movement, 259-267.
- Subjective color, 492-499.
- Tests, coördination, 592; human fitness, 590-611; reaction-time, 593; tapping, 592; tracing, 592.
- Textbook, abnormal, 746 f.; experimental psychology, 555 ff.; human relations, 202 f.; industrial, 739 ff., 751; industrial and business psychology, 392 f.; industrial psychology, 202; interviewing, 394; introductory psychology, 561 ff.; personality, 195 f., 203 f.; personnel administration, 200 f.; physiological acoustics, 735 ff.; psychiatry, 372; reading, 197 f.; sensation, 383 ff.; social, 752 f.; social psychology, 371 ff.; statistical methods, 566 f.; thinking, 194 f.; visual perception, 204.
- Temporal changes in memory of visual form, 647-658.
- Theory, behavior, 375 ff.; cochlear mechanics, 39-55; learning, 394; psychological, 386 ff.; visual form, 205-219.
- Thinking, 194 f.
- Threshold, form, visual, 205-219; recognition, visual function of word-frequency, 700-703; movement, visual, lower, 104-110.
- Tilt, effect on pursuit-movements, 475-483; effect on object orientation, 500-506.
- Time-order effects, theory of, 327-334.
- Traces, visual, 338-340.
- Transfer, rote-serial learning, 453-463.
- Uncertainty, methods for computing, 164-170.
- VEG scale, 677-683.
- Vexiersversuche, history of, 579.
- Vineland, social maturity scale, 754.
- Vision, acuity, Landolt's ring and fatigue, 509-512; binocular rivalry, 513-516; binocular subjective colors, 492-499; color, individual differences, 189 f.; color blindness, 182 ff.; critical flicker frequency, 632-646; critical review of theory, 182 ff.; 189 f.; discrimination (monkey), 517-520; fatigue, 509-512; form threshold, 205-219, 427-440; 'Maxwellian view,' 530-532; movement, thresholds of, 104-110; quantal fluctuations, 152 f.; red light, 727-730; reversible perspectives, 341 f.; rod function, 727-730; traces, 338-340.
- Vitamin deficiency, hoarding and, 124-128.
- Western Psychological Association, *see* Meetings.
- Word-frequency and learning, 138-140.

Tactual tasks, performance of, after brain injury, 220-240.







✓



